

欧米諸国におけるデータ分析、AIの活用による 会計検査・監査等の状況に関する調査研究

有限責任 あずさ監査法人 2022年3月

01 調査の概要

調査の概要

【目的】

本調査研究は、欧米諸国のデータ分析、AIの活用に関する調査研究、会計検査・監査の状況や実際の検査・監査事例について整理・分析することにより、日本の会計検査に有用な情報を提供することを目的として行った。

【対象】

- 本調査研究では、アメリカ、イギリス、オランダ、その他の国を対象に、文献調査及びインタビューにより以下の調査を実施する。
 - ▶ 欧米諸国の公的部門におけるデータ管理、データ連携、データ公開等の状況
 - ▶ 欧米諸国の会計検査院等におけるデータ分析、AIの活用等に関する調査研究、会計検査、監査等における事例
 - > まとめ

【期間】

• 2021年10月~2022年3月



調査プロセス

本調査研究では、以下の3つの段階を踏み調査を実施し、報告書を取りまとめた。

(1) 各国の公的部門 におけるデータ管理、 データ連携、データ公 開等の状況を整理



(2) データ分析・AI 政策に関する会計検 査等の状況の整理



(3) 調査結果の取り まとめ

- 調査対象は、アメリカ、 イギリス、オランダとした。また、欧州諸国等における取組も調査した。
- 国別に比較分析を することから、データ 管理、データ連携、 データ公開、データ利 用に関して共通の調 査すべきポイント※を 設けた。

第2章で報告

※「共通の調査すべきポイント」 は、「(3)の調査結果の取りまとめ」に記している内容である。

- 各国会計検査院の データベースでキー ワード検索を行い、 210事例を抽出した。
- 予め設定した抽出 基準でスクリーニング を行い、52事例を 抽出した(概要 表)。
- ・ 概要表に掲記され た報告事例から詳 細事例分析の対象 として22事例を抽 出した。
 - 第3章で報告

- 主な法律、公表物、所管官庁
- データをどのように集め、保存しているか。
- 公的機関でどのようにデータを連携しているか。会計 検査ではどのように活用しているか。
- データをどのように公開しているか。
- データをどのように利用しているか。データの利用の際に先端技術は使われているか。

第4章1節で報告

- データ収集: RPA・データマイニングにより情報収集、 データ整理、第三者機関からデータを購入した事例
- データ加工:前処理に多くの時間を要した事例 (個人情報、低い機械可読性)
- データ分析: クラスター分析・アソシエーション分析、 因子分析・主成分分析、回帰・決定木
- ビジュアル化
- AIを使った公共サービスを検査事例



02

各国のデータ分析、 AIに関する主要政 策の特徴



アメリカにおけるデータ分析、AIに関する政策の特徴

- アメリカでは、各連邦政府機関内の組織が高度なデータ分析やAIに関する研究開発を行っている。行政 管理局や監察総監評議会等が機関同士の連携を図り、情報技術研究開発の促進に取り組んでいる。
- アメリカの主な法律と公表物は以下のとおり。

図表 7 アメリカのデータ分析、AIに関する主要政策の変遷

	アメリカの主な法律	アメリカの主な公表物	日本の主な法律等(参考)
2000年以前	1996年:改正電子情報自由法		
2000年~ 2010年	2002年:電子政府(E-Government)法、 連邦情報セキュリティマネジメント法		 2001年:高度情報通信ネットワーク社会形成基本法(IT基本法) 2005年:e-文書法、個人情報保護法
	2009年:オープンガバメントに関する連邦指令		
2010年~ 2020年	2013年: デジタル政策 2014年: DATA法 (デジタル政策に基づく)、連邦情報セキュリティ近代化法 2015年: 不正削減およびデータ分析に関する法律 2016年: 連邦ソースコード政策 2018年: 政府データ公開法 2019年: 証拠に基づく政策立案法、AIにおけるアメリカのリーダーシップの維持に関する大統領命令	2010年:「連邦情報技術管理改革のための25の行動計画」 2012年:「デジタルガバメント戦略」、「21世紀の電子政府構築にあたって」 2018年:「技術評価: AI」(GAO) 2019年:「科学的機械学習の基礎研究ニーズに関するワークショップレポート: AIのコアテクノロジー」(労働省)	 2013年: 共通番号(マイナンバー) 制度関連法 2014年: サイバーセキュリティ基本 法 2016年: 官民データ基本法
2020年以降		2020年:「連邦データ戦略」、「AIアプリケーションの 規制のためのガイダンス」 2021年:「S&T:AIと機械学習の戦略的計画」 (国土安全保障省)、「技術評価設計ハンドブック」 (GAO)	・ 2020年: デジタル手続法、デジタル社会形成基本法・ 2021年: デジタル庁設置





アメリカにおけるデータ分析、AIに関する政策の特徴

データ管理

─ データをどのように集めているのか

•各連邦政府機関はデータ公開サイトで情報収集するか、省庁間や民間企業等と契約を交わして対象機関から情報収集している。

□ データをどのように保存しているのか

・連邦データ戦略、情報収集先との契約・合意条件、 各省庁内の管理計画、関連政策・法的要件を満た す方法でデータを保存している。

• 各連邦政府機関は、Data.govやITダッシュボードで 保有データを公開して誰もがデータにアクセスできるようにしている。GAO等が連邦政府機関に非公開データを要請した場合、連邦政府機関は対応しなければならない。

データ連携

□ 各省庁・各データベース間の連携をどのように図って いるのか

各連邦政府機関は各自のウェブサイトをData.gov、USA.gov、ITダッシュボード、USAspending.gov等の公開サイトと連携させている。この連携作業は、公開サイトへのログインIDを持つ連邦政府機関職員もしくは委託業者のみが実施できる。

□ 会計検査ではどのように活用しているか

• GAOは、受検機関からデータや関連資料を収集しているほか、統計情報や学術文献等はデータ公開サイトから収集している。また、必要であれば専門家等との会議を開催して情報収集している。





アメリカにおけるデータ分析、AIに関する政策の特徴

データ公開

| データをどのように公開しているのか

• 公共機関の保有データを公開するData.gov、給付金や税金等の情報を公開するUSA.gov、政府の支出データを公開するUSAspending.gov等のデータ公開サイトが整備されている。

─ 不正アクセス対応はどのようになっているのか

・データ公開サイトの管理機関が公開サイト利用者の 情報やアクセス状況を監視している。公開サイトへの 無断アップロードや公開情報の改ざんを防ぐために管 理機関は商用ソフトウェアを用いて監視している。

データ利用

データをどのように利用しているのか

• 連邦政府機関は、連邦データ戦略に従ってデータを利用している。GAO技術評価ハンドブックには、データ利用方法と留意点が事例とともにまとめられている。

データの利用の際に先端技術は使われているのか

• 保健福祉省OIGはニューラルネットワークとアルゴリズム を活用したデータ分析ツールを開発し、業績検査の監 査対象の特定、計画策定等の情報収集、財務検査 のターゲットの選定やデータ処理に利用している。デー タ分析にはクラウド技術やPython等も活用している。



GAOが作成した財務監査マニュアルにおいて、監査手続の性質、範囲、時期に関わる不正リスク対応時に用いる手 法の一例として、データマイニングやデータ抽出・分析ソフト(Interactive Data Extraction and Analysis)や有料ソ フト等の利用について記載されている。





イギリスにおけるデータ分析、AIに関する政策の特徴

- イギリスでは、デジタル・文化・メディア・スポーツ省、中央デジタル・データ室、政府デジタルサービス、公共部門情報局等の政府機関等が中心となり、公的・民間部門のデジタル化やデータ利活用を推進している。
- イギリスの主な法律と公表物は以下のとおり。

図表 12 イギリスにおけるデータ関連法、公表物

	イギリスの主な法律	イギリスの主な公表物	(参考) 日本の主な法律等
2000年以前	1998年:データ保護法(Data Protection Act 1998)		
2000年~ 2010年	2003年: PSI利活用に関するEU指令 2005年: PSI再利用規則(The Re-use of PSI Regulations 2005 (S.I.2005/1515))		2001年:高度情報通信ネットワーク社 会形成基本法(IT基本法) 2005年:e-文書法、個人情報保護法
2010年~ 2020年	2013年: PSI利活用に関するEU指令の改正 2015年: 公共データ再利用規則(Re-use of PSI Regulations 2015 (S.I.2015/1415)) 2016年: EU一般データ保護規則(The General Data Protection Regulation (2016/679)) 2017年: デジタル経済法(Digital Economy Act 2017) 2018年: データ保護法(Data Protection Act 2018)	2010年: Data.gov.uk運用開始 2011年: midataプロジェクト始動 2012年: GOV.UK開設 2017年: イギリスデジタル戦略(The UK Digital Strategy)公表	2013年 : 共通番号(マイナンバー)制度関連法 2014年 : サイバーセキュリティ基本法 2016年 : 官民データ基本法
2020年以降	2021年 : UK GDPR	2020年: 国家データ戦略(National Data Strategy)策定 2021年: 国家AI戦略(National AI Strategy)策定	2020年 : デジタル手続法、デジタル社 会形成基本法 2021年 : デジタル庁設置





イギリスにおけるデータ分析、AIに関する政策の特徴

データ管理

データをどのように集めているのか

各省庁は、国家データ戦略で定められているデータ品質フレームワークを各自で策定し、それに従って業務の実施に必要なデータを収集、管理している。

データをどのように保存しているのか

省庁が収集、所有するデータは、クラウド管理が義務付けられている。一方、各省庁が異なるクラウド上でデータを管理しているため、効果的なデータ連携には課題が残る。

─ データのアクセス管理をどのようにしているのか

国家サイバーセキュリティセンターがガイダンスを作成している。また、多くの公的・民間部門は国家統計局が掲げるデータ管理・アクセス原則を採用している。

データ連携

─ 各省庁・各データベース間の連携をどのように図って いるのか

- ・政府デジタルサービス局は各省庁が利用するデータ管理システムの統一を進めている。
- ・また、デジタル経済法で付与されている権限の活用 推進と、その他のデータ連携に対する法的な障壁の 解消の検討も進めている。

会計検査ではどのように活用しているか

- NAOは、受検機関からデータ・書類を収集し検査を 実施しているが、統計局等の第三者による公開デー タも必要に応じて用いている。
- NAOでは、これらの公開データを共通のフォーマットに標準化して集約し、主にVFM検査でのデータ分析の際に活用している。





イギリスにおけるデータ分析、AIに関する政策の特徴

データ公開

] データをどのように公開しているのか

• イギリスの公的部門が保有するデータは、政府デジタルサービス局が管理するdata.gov.ukで公開されている。

─ 不正アクセス対応はどのようになっているのか

・内閣府は行政サービスのデジタル移行に伴うセキュリティ確保、レガシーシステムやサポートが終了しているソフトウェアの入替え、システム、データ及びアクセス権限の更新等を行う予定である。

データ利用

データをどのように利用しているのか

公的部門におけるデータ利用は、データ諮問委員会 及びデータ・リーダーズ・ネットワークによって監督されている。

データの利用の際に先端技術は使われているのか

イギリス政府では、ビッグデータから得られる組織にとって有益なデータセットであるスマートデータの利用促進にも取り組んでいる。



- NAOのVFM検査では、ウェブスクレイピング等のデータ収集・分析手法や、データのビジュアル化が活用されている。
- NAOでは、財務検査にクラウド型の監査調書自動化システムや検査情報管理システムを導入している。また、リスク評価・計画ツールや試算表作成アプリケーションの開発及び活用等も行われている。
- さらに、NAOでは現在、今後の検査品質及び効率向上のために新たな検査手法を策定中である。この新検査手法 の適用をサポートするためのツールとして、目下ソフトウェアの開発も行っている。





オランダにおけるデータ分析、AIに関する政策の特徴

- オランダでは、AIや革新的技術を管轄する専門的な組織はなく、各行政機関にAI開発やデータ分析の専門部署を置いている。
- オランダの主な法律と公表物は以下のとおり。

図表 17 オランダのデータ分析、AIに関する主要政策の変遷

	四代 17 カランプのフ		
	オランダの主な法律	オランダの主な公表物	日本の主な法律等(参考)
2000年以前			
2000年~ 2010年	2000年:個人情報保護法(後にGDPRに置き 換わる) 2002年:司法データと刑事記録に関する法律		 2001年:高度情報通信 ネットワーク社会形成基本法 (IT基本法)
	2007年: 警察によって処理されたデータに関する 法律		・ 2005年:e-文書法、個人 情報保護法
2010年~ 2020年			・ 2013年: 共通番号(マイナンバー)制度関連法
	2015年:公的機関の情報の再利用に関する法律		・ 2014年 : サイバーセキュリ ティ基本法
	 2016年:オープンガバメント法 2017年:インテリジェンス及びセキュリティサービス に関する法律		・ 2016年: 官民データ基本 法
	2018年: 一般データ保護規則 (GDPR), ネット ワーク及びインフォメーションシステムのセキュリ ティーに関する法律	2019年:「AI戦略的行動計画(SAPAI)」、 「デジタル政府アジェンダ2019」	
2020年以 降	2021年:オープンガバメント法改正)	2020年:「オープンガバメント行動計(2020- 2022)」 2021年:「Understanding Algorithms」	・ 2020年: デジタル手続法、 デジタル社会形成基本法・ 2021年: デジタル庁設置



オランダにおけるデータ分析、AIに関する政策の特徴

データ管理

□ データをどのように集めているのか

オープンガバメント法に基づき、各行政機関は最善の 努力をもって積極的な開示をしている。

• PLOOI(オープンガバメント情報プラットフォーム)と 称する公的機関の公表資料を一元管理するプラット フォームを開発中である(施策版は公開中)。

─ データのアクセス管理をどのようにしているのか

 政府情報セキュリティ基準(Government Information Security Baseline)に基づき、官公 庁間での連携の改善、情報セキュリティの向上等が 図られている。

データ連携

- **各省庁・各データベース間の連携をどのように図って** いるのか
 - 政府のあらゆる文書が統一のプラットフォームPLOOIで標準化された形式で管理されているため、データの連携が容易となる。

・オランダ会計検査院は受検機関からの関連書類、 データの提出を求めているほか、必要に応じて、オラン ダ中央統計局のデータを用いて検査している。



オランダにおけるデータ分析、AIに関する政策の特徴

データ公開

• 公開されるデータには個人情報が含まれる可能性があるため、PLOOIでデータを公開する際、各種法規則に従い様々な公開ルールが定められている。

ナショナル・サイバー・セキュリティ・センター (NCSC)とデジタル・トラスト・センター (DTC) が、デジタル社会におけるセキュリティの維持を目的に連携している

データ利用

データをどのように利用しているのか

・オランダ中央統計局のアンケート調査(2018年)によると、回答のあった48%の行政機関がアルゴリズムを利用していると回答した。

データの利用の際に先端技術は使われているのか

・オランダ中央統計局のアンケート調査(2018年)によると、アルゴリズムを利用していると回答した行政機関のうち、機械学習やAIを用いて推定するケースベースのアルゴリズムの使用は47%であった。



オランダ会計検査院は、2021年に公表した「アルゴリズムを理解する」において、「アルゴリズムのための 検査フレームワーク」に基づく検査を実施した。



03

データ分析、AIに関する会計検査等の特徴



●「HHS-OIGの検査におけるAIの活用」(報告書p.24~)

- HHS-OIGは、他省庁と比較しても積極的にデータ利用やAIを用いたツール等の情報を公開している。
- HHS-OIGは、同省内の最高データ責任者室(Office of Chief Data Officers)と最高情報責任者室(Office of the Chief Information Officer)の協力を得て、クラウドコンピューティングとネットワークの最新化に取り組んでいる。

HHS-OIGデータ活用事例

- HHS-OIGは、アマゾン・ウェブ・ サービス (Amazon Web Service) クラウドとMicrosoft クラウドを利用。
- 統合データプラットフォームはアマ ゾン・ウェブ・サービス上で活用。
- 最新機能を備えた分析ツール を購入して利用することもあるが、 HHS-OIGは過去の不正事例 等から得られたデータを活用し て分析モデルやツールの開発を 行っている。

HHS-OIGのAI等の活用事例

- 一般的に、AIモデリングは調査担当者を利用対象として開発している。 様々なツールを活用することで高度な分析と機械学習を含むロジック ベースの手法を用いた監査が可能となっている。
- 同省内のOCDOのデータアナリスト、科学者、統計学者の協力を得て、 リスク特定のための高度なデータ分析とAIを活用した検査を実施してい る。
- 将来的には機械学習、自然言語処理、予測分析、データ分析とAI活用をHHS-OIGの調査及び監視ビジネスに統合することを目指している。 監査担当者が専門分野、専門知識、専門的判断が必要とされる作業に集中できるように、監査業務と管理業務におけるルーティン作業のRPAツールも構築している。例えば、クラウド上でAutomation Software Suites を活用して様々な業務を自動化している。





イギリスにおけるデータ分析、AIに関する会計検査等の特徴

●「NAOの監査の自動化」(報告書p.47~)

NAOでは検査におけるデータ分析の幅広い適用を推進しており、その推進の中心となるデータアナリティクスの専門家チームを二つ有している。各チームに所属するNAOの職員は、財務検査及びITに関する経歴を有しており、一部の職員は会計士資格も保有している。

- ①Analysis Hub(分析ハブ): データアナリティクスのためのアプリケーションの開発を行う10名規模のチーム。
- ②Financial Audit Data and Analytics Operations Team(財務検査データ・分析運用チーム): 4名体制で、検査チームが財務検査においてデータ分析を行う際、データの収集やクレンジング、AIMSへのデータのアップロード、及びアプリケーションの利用や結果の解釈支援といったサポートを行っている。
- なお、NAOによれば、これらのデータアナリティクスの利用については、現行の財務検査マニュアル (Financial Audit Manual) ではほとんど規定されていないとのことである。
- 他方、財務検査における仕訳帳の検査にはデータアナリティクスの利用を義務付けている。
- NAOでは、財務検査を担当する職員に対して、主にデータ分析アプリケーションの利用方法に焦点を当てた研修を行っている。また、さらにデータアナリティクスについて知識を深めたいという財務検査担当職員に対しても、アプリケーションの開発に関する研修等が提供されている。
- NAOへのインタビューによれば、現在データ分析の活用に関するパフォーマンス指標として、データアナリティクスツールを活用した年間検査件数を用いている。
- 約400の被検査機関への検査のうち、ツールを用いて検査を行った検査チーム数についても、パフォーマンス目標が設定されている。





イギリスにおけるデータ分析、AIに関する会計検査等の特徴

●「NAOの監査の自動化」(報告書p.43~)

一貫性のある、より品質の高い検査を行えるよう、サンプリングや分析等の検査手続の一部を自動化する複数のアプリケーションをNAO内部のチームで開発

総勘定元帳と試算表の調整や試算表と財務諸表の調整を自動化する試算表作成アプリケーション

突合作業を数秒で行い、金額の不一致があればそれを特定することができるもので、手作業で行うよりもミスが発生しづらく、現在では財務検査で広く使われている。

グループの連結試算表を作成するためのアプリケーション

検査対象となる機関の中には多数の事業体を有する機関もあるが、それらの事業体から提出された書類すべてを用いてグループの連結試算表を作成する。各書類を連結試算表に統合するプロセスの自動化を可能にするとともに、データを整理することでグループ全体の分析を詳細にかつ容易に掘り下げて実施することができるようになった。

プロセスに特化したデータ分析出力ツール

Purchase-to-pay(購買管理システム)やPayroll (給与計算システム)といった主要なビジネスプロセスの検査に特化したデータ分析出力ツールの開発も行っており、 幾つかの機関に対する検査で運用している。

Confirmation.comによるオンラインサービス

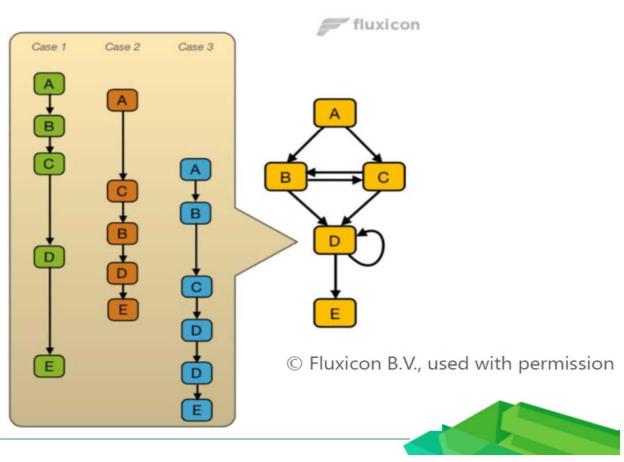
2017年からConfirmation.comという組織が提供するオンライン銀行残高確認サービスも利用している。NAOによれば、通常の検査手続の一環として検査対象機関が使用している銀行の残高を確認する必要があるが、Confirmation.comによるオンラインサービスを利用することにより銀行から確認書を入手するプロセスが簡素化かつ迅速化されている。



●「ヨーロッパ会計検査院のプロセスマイニングの活用」(報告書p.77~)

What is process mining?

Case	Activity
1	А
2	A
1	В
1	С
3	Α
2	С
3	В
2	В
1	D
2	D
2	E
3	С
3	D
1	E
3	D
3	E



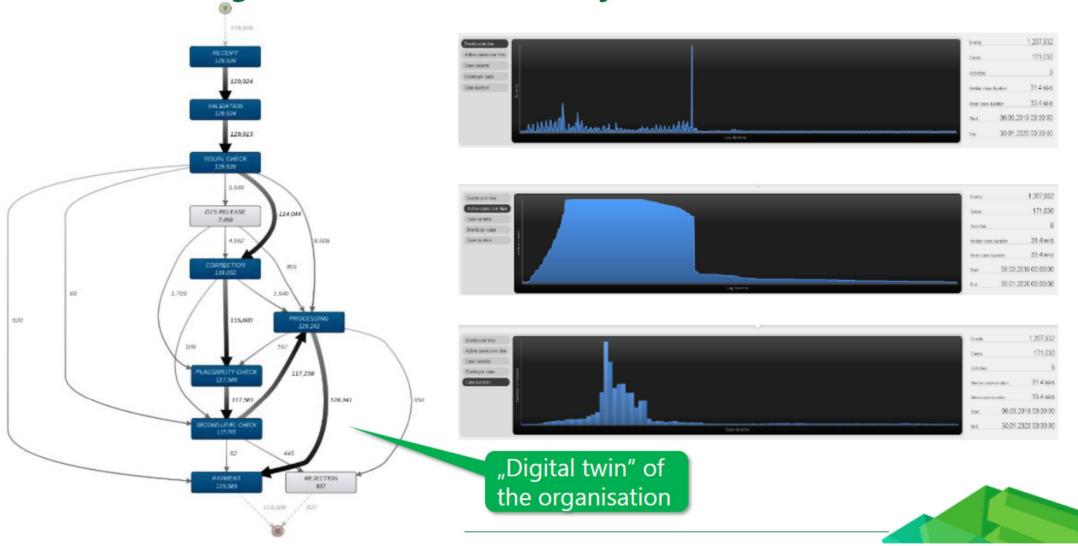


出所: Zsolt VARGA, ECALab 「Process Mining in EU audit: a case study」

https://eurosai-it.org/images/public-documents/meetings/6 ECA-Zsolt-Varga Process mining EUROSAI.pdf



Process mining as a tool for visual analytics

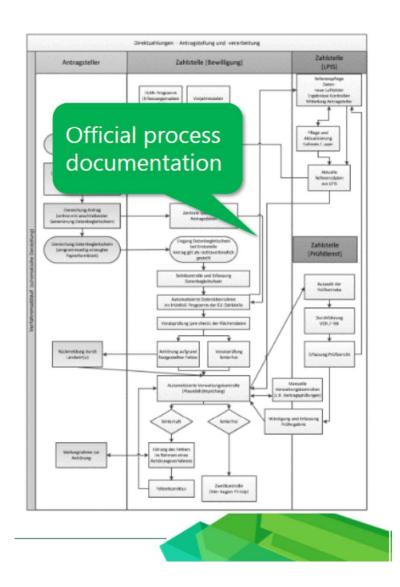


出所: Zsolt VARGA, ECALab 「Process Mining in EU audit: a case study」

https://eurosai-it.org/images/public-documents/meetings/6 ECA-Zsolt-Varga Process mining EUROSAI.pdf



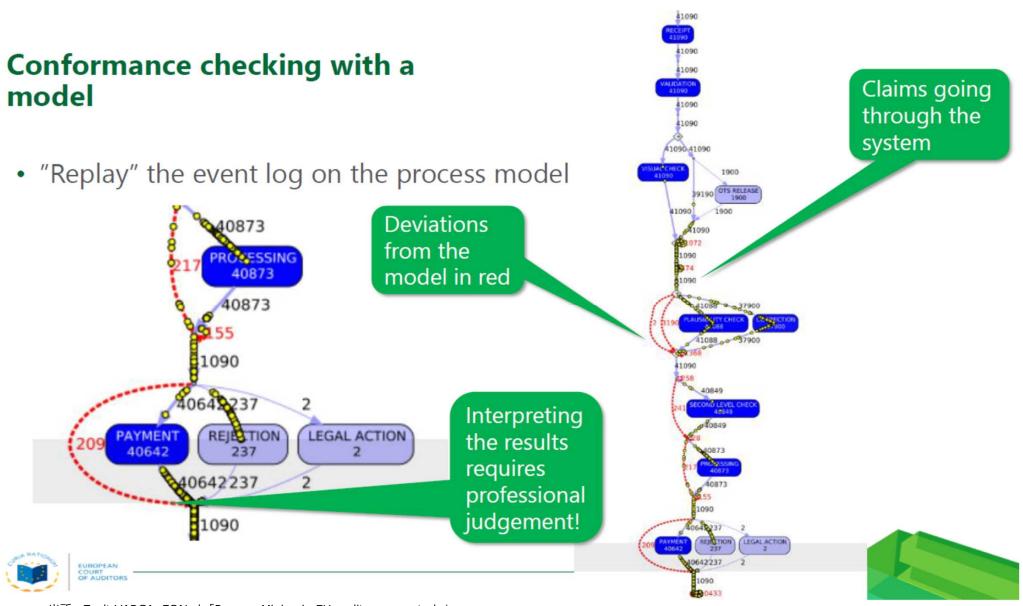
Comparing process maps 41,090 129,924 41,089 Single All claims claims only 120 013 41.081 MSUAL CHECK 41,090 17,724 OTS RELEASE CORRECTION 41 000 20.018 2,390 520 115,600 OTS RELEASE 28.361 2,394 2,137 528.341 PLAUSBILITY CHECK 45,615 41,013



出所: Zsolt VARGA, ECALab 「Process Mining in EU audit: a case study」

https://eurosai-it.org/images/public-documents/meetings/6 ECA-Zsolt-Varga Process mining EUROSAI.pdf





出所: Zsolt VARGA, ECALab 「Process Mining in EU audit: a case study」
https://eurosai-it.org/images/public-documents/meetings/6 ECA-Zsolt-Varga Process mining EUROSAI.pdf



- プロセスマイニングによって、単一アプリケーションの8%が抜き打ち調査の対象となっていることが明らかになった。他方で、5月15日の締め切り後に受け取った単一アプリケーションの抜き打ち調査率は0.7%であった(5%の規定未満である)。
- ECAが確認したところ、一部のグループにだけ抜き出して5%の規定を 遵守する法的要件はなく、全体で5%の規定が守られていれば良いと のこと。



プロセスマイニングによって、不正な取引の発見が可能になる。

※上の例では、結果としてはルールの範囲内



●「ヨーロッパ会計検査院のブロックチェーンの活用」(報告書p.81~)

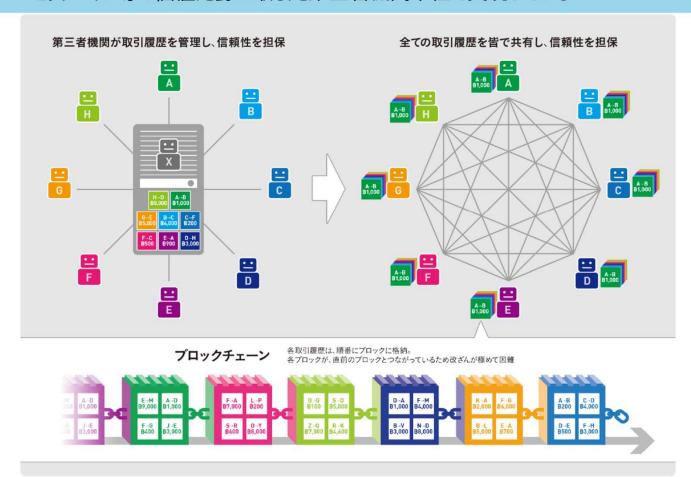
ヨーロッパ会計検査院におけるブロックチェーンの活用状況

- ヨーロッパのブロックチェーンインフラストラクチャー(EBSI)は現在、EUのほぼ全域をカバーしており、14の加盟国が24ノードを運用しており、欧州委員会が6ノードを運用している。
- ECAレジストリ(検査書類の利用)が最初の利用例として選ばれた。
- ECAレジストリは、ヨーロッパのブロックチェーンノードネットワークによって支えられている。
- レジストリは文書のストレージとしても機能するため、さまざまな利害関係者が文書を アップロードしてブロックチェーンに登録することができ、同じWebインターフェイスを使用 して文書を比較又は認証できる。



ブロックチェーンとは

• ビットコイン等の価値記録の取引を第三者機関不在で実現している

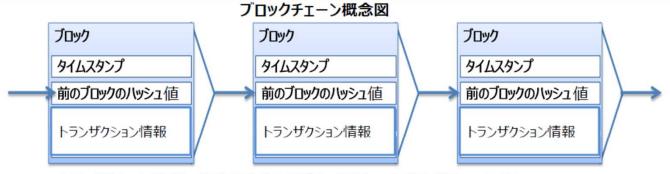


出所:商務情報政策局情報経済課(2016)「平成27年度我が国経済社会の情報化・サービス化に係る基盤整備(ブロックチェーン技術を利用したサービスに関する国内外動向調査)報告書概要資料」

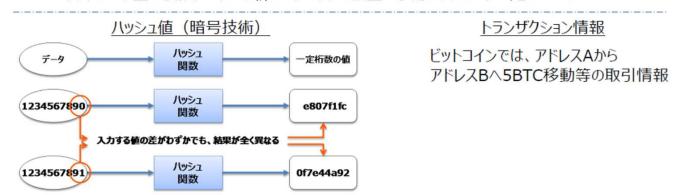


ブロックチェーン技術とは

- ビットコインを実現させるために生まれた技術であり、いくつかの暗号技術がベース
- P2Pネットワークを利用してブロックチェーンデータを共有し、中央管理者を必要と せずにシステムを維持することを実現



- ・トランザクション情報の集合等を含んだブロックがチェーン状に連なっているもの
- ・ネットワーク上の複数ノードが、新しいブロックを相互に承認し、チェーンに足していく



出所:商務情報政策局情報経済課(2016)「平成27年度我が国経済社会の情報化・サービス化に係る基盤整備(ブロックチェーン技術を利用したサービスに関する国内外動向調査)報告書概要資料」



検査にブロックチェーンを活用することのメリット

- プロセスマイニングにブロックチェーンを用いることは、検査のアウトプットの処理だけでなく、検査の 処理手順の点でも意味がある。
 - ▶ 今後は、文書だけでなく、データベースからのデータ抽出に対しても適切な検査証跡を確保することが 重要になってくる。
 - データベースのエクスポートは非常に大きくなる可能性があるが、それらのデジタル証跡(デジタル指 紋)は小さく、ブロックチェーンレジストリにより管理できる。
- 検査証跡のデータの整合性を確保するために、高価な商用ソフトウェアで再認定されたアプリケーションを使用する必要はない。データの処理と分析に使用したコンピュータースクリプトを分析結果の画像と最終レポートとともに証明することで、完全な透明性を担保できる。
- 入力、中間ステップ、及び出力を証明することができ、検査の再実行が可能になる。また、ブロックチェーンをプロセスマイニングに適用することで、オリジナルのファイルと検査対象機関から提出したファイルとの整合性を確認することなどができる。



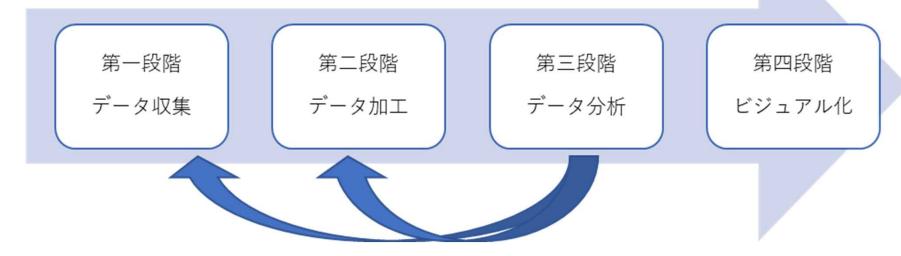
ブロックチェーンを活用することで、より安全かつ信頼性の高い検査が可能になる。



データ分析、AIに関する会計検査等の特徴

●「データ分析概論」(報告書p.82~)

データ分析では、より得たい結果を得るため、第一段階から第三段 階までのサイクルを繰り返し、データと結果を精緻化するのが通常





データ分析、AIに関する会計検査等の特徴

●「データ分析概論」(報告書p.82~)

図表 32 目的別のデータ分析手法

目的	手法	分類/回帰	機械学習に おける扱い
関連性の分析	相関分析	_	—
グループ化	クラスタリング		教師なし学習
次元の削除	因子分析、主成分分析		教師なし学習
現象の説明、要因の分析、予測	線形回帰 ロジスティック回帰 決定木	回帰 分類 分類/回帰	教師あり学習
予測	ランダムフォレスト ニューラルネットワーク	分類/回帰 分類/回帰	教師あり学習

出所:有賀・大橋(2019)を基にあずさ監査法人が作成





データ収集、データ加工において参考になる事例

- ●「Supplemental Nutrition Assistance Program: Disseminating Information on Successful Use of Data Analytics Could Help States Manage Fraud Risks(補充的栄養支援プログラム:データ分析に係る成功事例に関する情報を広めることは、州が不正リスクを管理するのに役立つ可能性がある)」【2018年】(報告書p.92~)
- 補充的栄養支援プログラム(SNAP)に係る給付金の不正受給や不正使用のリスクを低減するために、各州はデータマッチングやデータマイニング等のデータ分析を用いてSNAPを用いた不正購買パターンや傾向を特定している。
- · GAOは、そのような不正行為に対抗するための州の取組について検査した。





図表 37 州のSNAP機関がデータマッチングとデータマイニングを利用して 不正を検出する方法の例

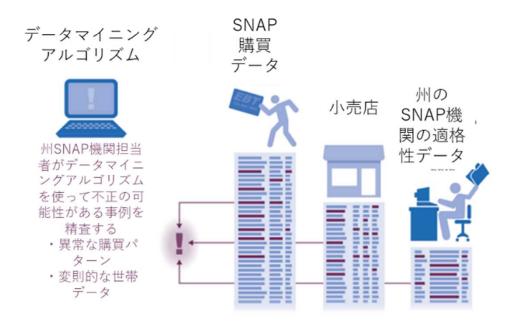
データマッチング

様々な目的で収集された記録とファイル 等の情報を比較し、不正の可能性を意味 するエラーや不正な情報を検知すること

州SNAP機関担当者が 機関 担当者 州SNAP機関担当者が 購買記録に含まれる データを二次データと マッチングさせて顧客 情報や未報告事項を検 出する

データマイニング

大量のデータの中から不正パターン、 相関関係、異常等を検知するためにコン ピュータアルゴリズムを利用すること



出所:検査報告書(GAO-19-115)





①検査方法 🗒

- コロンビア特別区とノースダコタ州、ワシントン州のSNAP受給世帯の取引データの分析。
- 七つの州(コロンビア特別区、マサチューセッツ州、ミシシッピ州、ニューメキシコ州、ノースダコタ州、 ワシントン州、ウィスコンシン州)のSNAP機関の担当者に対してインタビューを実施。
- USDA食品・栄養局が州に授与した助成金の文書を確認。

②検査による発見





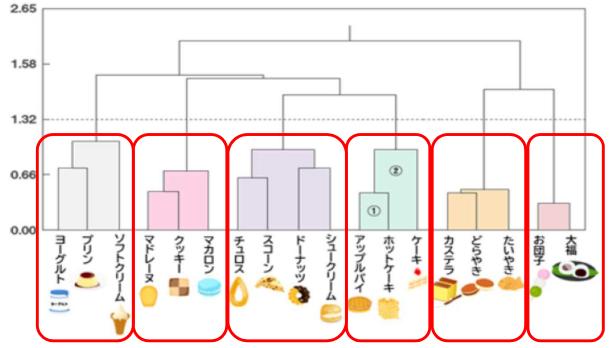
- GAOが選択した3州における2017年度のSNAPデータを分析したところ、全体で約2%の世帯が、 州内及び州外で不正な購買を行っていたことが判明した。
- GAOが調査した7州すべての関係者は、データ分析に当たりデータマッチングを行ったと述べている。 このうち5州の担当者は、データマイニングを含むより高度なデータ分析を行うことにより、潜在的な 不正の特定に役立つと述べている。
- また、より高度なデータ分析を用いることにより、自動的に不正を検出できるようになるだけでなく、 調査対象事例に優先順位をつけることで、不正の可能性が最も高い事例に検査資源を集中させることができるという利点を挙げている。

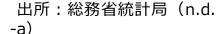




データ分析において参考になる事例

- ●「Unemployment Insurance: States' Reductions in Maximum Benefit Durations Have Implications for Federal Costs (失業保険:州による最長給付期間の短縮は連邦費用に影響を与える) 【2015年】 (報告書p.117~)
- 各州が失業給付の最長期間を短縮した状況や、短縮が個人に与える影響、連邦政府の失業給付に係るコストへの影響等を分析した。
- GAOは、データの分析に際してクラスター分析を利用した。









50州すべてとコロンビア特別区のうち、同じカテゴリーに分類される州(給付期間短縮州)

実際に給付期間を短縮した州

① アーカンソー、フロリダ、 ジョージア、ミズーリ、ノース カロライナ、サウスカロライナ ②イリノイ、 カンザス、ミシガン 実際に給付期間を 短縮しなかった州

インディアナ、オハイオ、 テネシー

特性が似ている2つのグループを特定

各クラスターに属する 州を比較することで、 同様の状況にある 州が給付期間を短 縮しないことを選択 した理由を調べるこ とができた。



期間を短縮した九つの州について分析を行った結果、期間を短縮した州は、しなかった州と比較して以下のような点で他の州とは異なっていることが分かった。

- ✓ 不況前の信託基金のバランスが悪い
- ✓ 課税対象となる資金の総額が少ない
- ✓ 連邦政府からの借入れが多い
- ✓ 失業率が高い
- ✓ 労働組合への参加率が低い
- ✓ (州議会と州知事の)政治的な同質性が高い



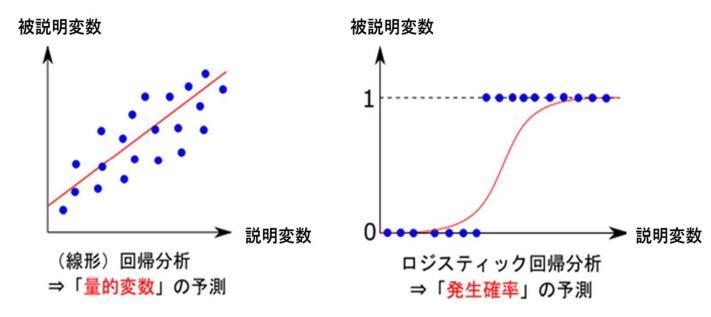


イギリスにおけるデータ分析、AIに関する会計検査等の特徴

データ分析において参考になる事例

- ●「Staying the course: the retention of students in higher education (最後までやり遂げる:高等教育における学生の定着率) 【2007年】 (報告書p.207~)
- ・ 学生の定着率改善に対する高等教育部門による取組が向上しているか検査を行った。
- ・検査の一環として、高等教育統計局から得た学生データを用いた回帰分析(線形回帰分析及びロジスティック回帰分析)を行い、学生の定着率に影響する要素について検討した。

図表 44 ロジスティック回帰分析のイメージ図



出所:知識のサラダボウル(n.d.)(左記出所を参考にあずさ監査法人が作成。)





イギリスにおけるデータ分析、AIに関する会計検査等の特徴

○線形回帰分析の結果

- NAOでは、121の教育機関が提供した個別学生記録データを用いて線形回帰分析を行い、特定の社会的グループからの学生の割合やAレベルのスコア等の複数の説明変数が変化した場合における教育機関ごとの継続率(被説明変数)への影響を分析した。
- 決定係数R²からは、NAOが作成した回帰モデルはフルタイムの学生の継続率の変動を71%説明していることが示されており、NAOはこのモデルは比較的良いモデルとしている。
- 一方、パートタイムの学生を対象に作成した回帰モデルは、継続率の変動を21%のみしか説明できていないことが分かったため、検査報告書ではパートタイムの学生に関する回帰モデルは使用されていない。

図表 82 継続率に影響を及ぼす変数

変数	回帰係数	p値
医学・歯学を専攻している学生の割 合	0.134	0.0001
教育学を専攻している学生の割合	0.131	0.0000
医学関連科目を専攻している学生の割合	0.035	0.0389
クリエイティブアーツ (創造的芸術)・ デザインを専攻している学生の割合	0.033	0.0003
高等教育への就学率が平均的な地域	0.029	0.0001
Aレベルの平均スコア	0.006	0.0000
ヨーロッパの言語、文学及び関連科目 を専攻している学生の割合	-0.060	0.0068
21歳以上の学生の割合	-0.070	0.0021
定数項	0.785	0.0000

出所:検査報告書(HC616)





イギリスにおけるデータ分析、AIに関する会計検査等の特徴

○ロジスティック回帰分析の結果

- 学生の継続率に可能性を与える要因のさらなる理解を目的に、NAOではフルタイムの学生 299,000人及びパートタイムの学生229,000人のデータを用いて、フルタイムとパートタイム別 の二つのモデルを構築し、学生の年齢や学習科目等の説明変数と、学生の学習継続(被説 明変数)との関係を分析した。
- なお、高等教育課程の継続に当たっての学生の選択は「継続する」又は「継続しない」の2値の どちらかとなるため、ロジスティック回帰分析を使用した。



- 学生の継続率に最も影響を及ぼす要素は、その学生がフルタイムで勉強しているか、又はパートタイムで勉強しているかである(オッズ比3.3)。フルタイムの学生の方が、パートタイムの学生よりも継続率が高くなる可能性がある。
- フルタイムの学生の場合、学生の継続率に最も大きな影響を及ぼす要素は、入学時の資格 (Aレベルのスコア)である。
- パートタイムの学生の場合、特にファウンデーション学位を過去に取得しているパートタイムの学生の場合は、学士課程に継続して在籍する可能性が高まる。 等が明らかになった。





イギリスにおけるデータ分析、AIに関する会計検査等の特徴

データ収集、ビジュアル化において参考になる事例

- ●「Public service markets: Putting things right when they go wrong (公共サービス市場:失敗を是正す
- る) 【2015年】(報告書p.227~)
- イギリス政府では、公共サービスの利用に際して問題が生じた際に、利用者が苦情申立及び救済申請を行えるようにしており、そこで得られた苦情に関するデータは公共サービスの改善等に活用されることとしている。NAOでは、この苦情・救済措置についてVFM検査を実施した。
- 検査方法の一環として、NAOではR言語と Pythonを用いて、ソーシャルメディアの分析を 行った。そのうち、保護者向けのウェブサイトに投 稿された内容の分析結果については、ワードクラ ウドにより視覚化して報告した。
- NAOはオンブズマンに対して、ソーシャルメディア の分析を活用した課題の特定を検討するよう勧 告した。



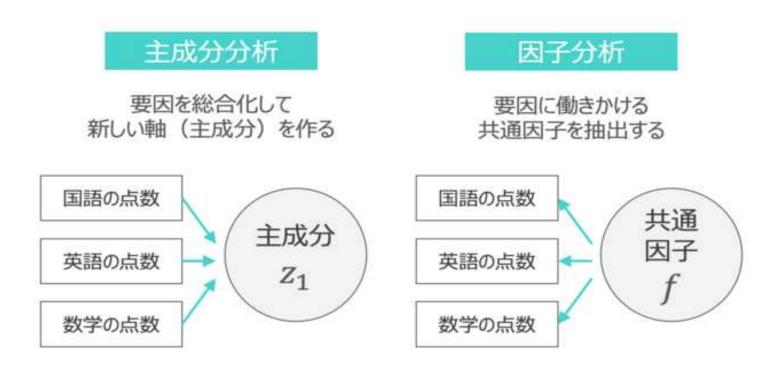
出所:検査報告書(HC84)





- 因子分析は、多変量解析の手法の一つ。多変量データに潜む共通因子を探り出すための手法で、教師なし学習に分類される。
- モデル式の形状等から主成分分析と混同されることもあるが、主成分分析は観測データから合成スコアを構築することが目的であるのに対し、因子分析は観測データが合成量であると仮定し、個々の構成要素を得ようとすることが目的であり、両者は因果関係が逆になっている。

図表 55 因子分析と主成分分析のイメージ



出所: Marketing Research Journal (2021)





データ分析において参考になる事例

- ●「Public Health: A Health Status Indicator for Targeting Federal Aid to States (公衆衛生:連邦による州への支援対象となる健康状態の指標) 【1996年】(報告書p.161~)
- 議会はGAOに連邦公衆衛生補助金の配分の目標設定に使えるような、州の人々の健康 状態を表す指標を特定するよう求めた。
- GAOは、「ヘルシーピープル2000」指標やリライアスター社が開発した複合健康状態指標について主成分分析を行い、その主成分(ヘルシーピープル2000指標の全体的な傾向を示す)と若年死亡率、貧困者数との関係をステップワイズ回帰によって分析し、若年死亡率と主成分との相関、すなわち、若年死亡率と「ヘルシーピープル 2000」指標全体やリライアスター指標全体との相関が高いことを示した。
- 若年死亡率が、各州の様々な健康状態を相当程度説明できていることを示し、若年死亡 率を連邦政府の公衆衛生補助金の配分式に含めることが妥当であるということを示した。





図表 65 主成分分析の結果

主成分	寄与率	累積寄与率	寄与度の比率 (寄与率の合計 を100)
1	42.1	42.1	55
2	17.4	59.5	23
3	10.7	70.2	14
4	6.3	76.5	8

出所:検査報告書(GAO/HEHS-97-13)

図表 66 各主成分のヘルシーピープル 2000 指標の 各指標に対する重み

ヘルシーピープル 2000指標	1 (42%)	2 (17%)	3 (11%)	4 (6%)
死亡率指標(Mortality indicators)				
総死亡率(Total mortality)	.9264	0553	1459	.2557
心血管関連死(Cardiovascular deaths)	.8802	1506	3086	.0626
肺がん死(Lung cancer deaths)	.6476	2370	3921	.3749
乳がん死(Breast cancer deaths)	.0447	5010	4876	.2801
自動車事故死(Motor vehicle deaths)	.5644	.7379	0742	.1008
自殺(Suicides)	0116	.7414	.3660	.4271
殺人(Homicides)	.8646	0970	.3097	0020
乳児死亡率(Infant mortality)	.7819	.0662	3049	1345
労災による死亡(Work injury-related deaths)	.0182	.6458	0688	0482
疾病発生率の指標 (Disease incidence indicators)				
AIDS	.2441	7371	.4661	0459
梅毒(Syphilis)	.8241	.0194	1338	1413
結核(Tuberculosis)	.5072	4229	.5359	2064
リスク要因(Risk factors)				
劣悪な大気質(Poor air quality)	0779	3571	.4027	.6837
産前産後ケア(Prenatal care)	.5331	.3660	.5341	.0038
子供の貧困(Childhood poverty)	.7737	0606	.0288	1107
低出生体重(Low birth weight)	.8538	1136	.0940	1468
若年出産(Births to adolescents)	.8969	.3067	.0773	.0034

※0.6以上は太字

出所:検査報告書(GAO/HEHS-97-13)





- 第1主成分は、ヘルシーピープル 2000指標のうち九つと、他の選択された指標のうち二つ(若年死亡率と貧困者数)と0.60以上の相関があった(主成分係数が0.60以上)。これらのうち八つの指標との相関は0.80を超えていた)。
- **若年死亡率は、第1主成分との相関が最も高かった**ため、ステップワイズ回帰法により、モデルに含めるべき最初の変数として選択された。若年死亡率だけで、この成分の変動の86%を説明することができた。心血管関連死(cardiovascular-related deaths)を加えると、説明される変動がさらに 7 %増加し、決定係数R2 は 93%になった。若年出産(Births to adolescents)を加えると、説明される変動は4%増加し、殺人(homicide)と低出生体重(low birth weight)によって、説明される変動はそれぞれ1%増加した。
- 第1主成分と、ヘルシーピープル 2000指標との相関係数は主成分係数として表に整理されているが、第1主成分と若年死亡率と貧困者数との関係は表で示されていないため、非常に分かりにくくなっている。

第1主成分が、ヘルシーピープル 2000指標のうち 九つと相関高い = ヘルシーピープル 2000 指標の様々な指標の動き を説明している。

若年死亡率は、 第1主成分との 相関が最も高く、 変動の86%を説明 若年死亡率を指標として使うだけでも、各州の全体的な健康状態を示せると言える。





データ分析において参考になる事例

- 「Effect of Regulations on Small Business Lending and Institutions Appears Modest, but Lending Data Could Be Improved (コミュニティバンク 小規模企業向け融資と金融機関に対する規制の影響は軽微だが、融資データの改善が望まれる」) 」【2018年】 (報告書p.124~)
 - 2007年~2009年の金融危機以降、規制当局が規制を大きく変えた。議会の求めにより、GAOは、2010年以降の規制変更がコミュニティ・バンクと中小企業向け融資に与えた影響を評価した。
 - ・ 2010年以前(2003年から2009年まで)のマクロ経済や地域市場の状況、及び測定可能な要因である銀行の特性に関するデータを用いて、ロジスティック回帰を行い、予測モデルを推定した。その予測モデルを用いた推計結果と、2010年以降の実際の数値との差を評価し、規制の影響を評価した。
 - ※規制の影響がなければ、予測モデルを用いた推計結果と、実際の数値との乖離が小さくなる。





Y imt = B0 + MACRO t B1 + LDEMOG mt B2 + LCOMP mt B3 + BANKCHAR it B4 + e imt

①合併モデル

t年において、コミュニティ・バンクが買収された場合は、Yを1、されていない場合は0とした。



ロジスティッ ク回帰

②新規銀行設立モデル

・ 新しいコミュニティ・バンクが、t年のm市場で設立されたかどうかをモデル化した。コミュニティ・バンクが新たに設立された場合はYを1、設立されなかった場合はOとした。



③中小企業向け融資モデル

コミュニティ・バンクiによるt年の小企業向け融資をモデル化した(Yを融資額とした)。

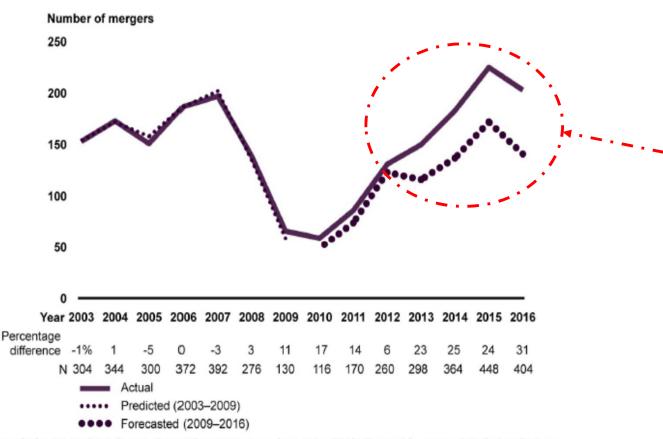
④税引前資産利益率モデル

・ コミュニティ・バンクiのt年の税引前資産利益率をモデル化した(Yを税引前資産利 益率とした)。





図表 47 合併モデル



予測と差があり、 規制環境の変化 が影響している可 能性が否めない。

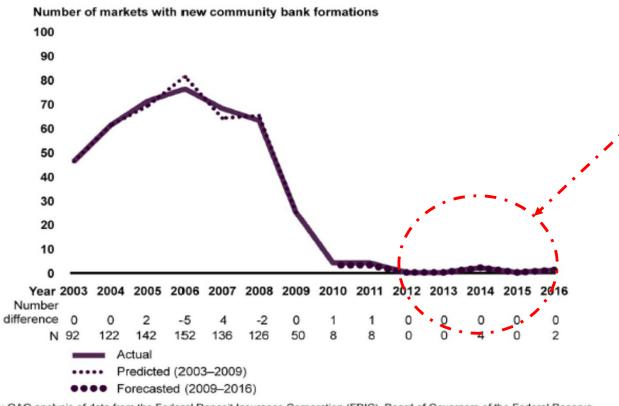
Source: GAO analysis of data from the Federal Deposit Insurance Corporation (FDIC), Board of Governors of the Federal Reserve System, Bureau of Economic Analysis, Bureau of Labor Statistics, Census Bureau, Federal Housing Finance Agency, and National Credit Union Administration. | GAO-18-312

出所:検査報告書(GAO-18-312)





図表 48 新規銀行設立モデル



予測と差がなく、規制環境の変化の影響は少ない。

Source: GAO analysis of data from the Federal Deposit Insurance Corporation (FDIC), Board of Governors of the Federal Reserve System, Bureau of Economic Analysis, Bureau of Labor Statistics, Census Bureau, Federal Housing Finance Agency, and National Credit Union Administration. | GAO-18-312

出所:検査報告書(GAO-18-312)



04まとめ

調査結果のポイント

- ▼ アメリカ、イギリスでは、検査におけるデータ分析に活用が進んでいる。
- 特に、アメリカでは、HHS-OIGにおいては、同省内のOCDOのデータアナリスト、科学者、統計学者の協力を得て、リスク特定のための高度なデータ分析とAIを活用した検査を実施している。将来的には機械学習、自然言語処理、予測分析、データ分析とAI活用をHHS-OIGの調査及び監視ビジネスに統合することを目指している。
- ECA Labは検査にプロセスマイニングを活用したり、ブロックチェーンの活用をしたりする取り 組みを進めている。
- アメリカGAOやイギリスNAOの検査報告書には、ロジスティック回帰や主成分分析等の統計分析の技術を活用した事例が見られる。
- また、NAOでは、ウェブスクレイピングやテキスト分析、センチメント分析(ビジュアル化)なども活用している。



本資料は、会計検査院の委託により有限責任 あずさ監査法人が実施した調査結果を取りまとめたものです。本資料は委託者である会計検査院に対して提出したものであり、本資料を閲覧あるいは本報告書のコピーを入手閲覧した第三者の本資料の利用により被る損害・損失に対して、有限責任 あずさ監査法人は直接ないしは間接の責任を負うものではありません。有限責任 あずさ監査法人は、調査時点で入手した情報に基づき本報告書を適時に取りまとめるよう努めておりますが、本資料の内容は、本調査の対象に含まれない特定の個人や組織が置かれている状況への適用を想定しておりません。また、本資料の情報を受け取った時点及びそれ以降において、その情報が最新かつ正確、完全であることを保証するものではありません。

