

行政 DX と自治体業務の効率化に向けた指標 及びデータ活用スキーム研究

ログを用いた低負荷・継続的なデジタル測定による
職員・住民の負担軽減

2026 年 07 月 31 日

一般財団法人 GovTech 東京

0. はじめに	6
0.1 調査の目的	6
0.2 背景：行政 DX の現状と課題	6
0.3 行政 DX の効果検証が困難な理由	7
0.3.1 指標の未整備	8
0.3.2 データ不足	8
0.3.3 前後比較が困難	8
0.3.4 自治体間比較が困難	9
0.3.5 行政 DX の対象領域の異質性（G2C／G2B／G2E／G2G）	9
0.4 調査スコープ	11
1.行政 DX における効果検証の枠組み	13
1.1 指標の概要	13
1.1.1 指標の定義と範囲	13
1.1.2 指標の分類	13
1.1.3 ステークホルダー別指標	14
1.2 測定方法の概要	15
1.2.1 実測	15
1.2.2 試算	16
1.2.3 ログデータ分析	16
1.2.4 ワークサンプリング法	16
1.2.5 アンケート	16
1.2.6 自己申告法	16
1.3 指標×測定法による事例の整理	17
1.3.1 RPA（Robotic Process Automation）	17
1.3.2 地理空間情報の活用	18
1.3.3 窓口業務改善	18
1.3.4 AI 活用	19
1.4 評価フレームワーク	20
1.4.1 事前評価（ex-ante）と事後評価（ex-post）の位置づけ	20
1.4.2 ロジックモデル	21
1.4.3 バランス・スコアカード	21
1.4.4 公共価値の三角形（Moore's Triangle）	21
1.4.5 PICO	22

1.5 分析方法	22
1.6 現行枠組みの課題	25
1.6.1 指標未整備と不統一	25
1.6.2 アウトカム不足	26
1.6.3 測定コスト・運用制約	26
1.7 小括：効果検証枠組みの整理	27
2.自治体業務と行政 DX の現状	29
2.1 自治体業務の構造と特徴	29
2.1.1 業務分類（コア／非コア）	29
2.1.2 プロセス分類（法令／住民／内部／更新）	31
2.2 DX 取り組み状況と課題	32
2.2.1 標準化・共通化の進展	32
2.2.2 DX 進行状況	32
2.2.3 東京都の事例	32
2.3 効果検証が困難となる要因	33
2.3.1 業務量の可視化不足	33
2.3.2 成果の定量化が困難	34
2.3.3 PDCA 不全（特に G2E・G2G 領域）	34
2.3.4 予算執行と人事ローテーション	35
2.4 小括	35
3. 「測定負担」を増やさないデジタル測定の考え方	37
3.1 なぜ従来の測定は負担を生み、継続しにくかったのか	37
3.1.1 計測が生む現場への負担	37
3.1.2 効果検証に耐えうる観測時点の不足	38
3.1.3 長期的測定の難しさ	38
3.2 デジタル測定としてのログデータ	39
3.2.1 ログデータにより把握可能な成果の範囲	41
3.2.2 ログが実現する自動計測	41
3.2.3 全件・連続データとしてのログの特性	42
3.2.4 測定行為を伴わないデータ取得がもたらす利点	43
3.2.5 デジタル測定が内在するリスク	43
3.3 ログによって可能になる効果検証の拡張	44
3.3.1 前後、自治体間の反実仮想分析	45
3.3.2 工程分析・ボトルネック可視化	45

3.4 デジタル測定を活用した評価の実例	46
3.4.1 群馬県前橋市・高崎市・伊勢崎市	47
3.4.2 カナダ政府	47
3.4.3 静岡県裾野市	49
3.4.4 東京都練馬区	50
3.4.5 民間事例	51
3.5 小括	51
4. ログを活用し、業務と測定を一体化する行政 DX 効果検証モデル	
の検討	53
4.1 モデル設計の基本方針	53
4.1.1 観測データと非観測データの整理	54
4.1.2 業務ログを中心とした「ログを用いた測定体系」	55
4.1.3 測定負担を最小化し、継続運用可能な仕組み	55
4.2 業務ログに基づく処理時間・滞留状況の把握モデル	56
4.2.1 業務単位の処理時間・滞留状況の把握	56
4.2.2 ボトルネック特定手法（待ち時間／操作回数）	56
4.3 導入ステップ（運用モデル）	56
4.3.1 ログ収集の最低要件	57
4.3.2 データ収集 → 分析 → 改善の年次サイクル	58
4.3.3 小さく始めて継続するための運用ガイドライン	58
4.4 デジタル測定モデルの活用ケース	59
4.4.1 G2C：住民手続の負担軽減の可視化	59
4.4.2 G2E：職員業務の工程分析	60
4.4.3 G2G：連携業務の遅延箇所の特定	60
4.5 ログ観測と補完手法の接続	60
4.5.1 ログ指標とアウトカムを接続する仮説設計	60
4.5.2 補完手法の実施要件	61
4.6 小括：デジタル時代の行政評価はデジタルで測るモデルへ	62
5. まとめ	64
5.1 本研究の提言	64
5.2 本研究の限界と課題	66
5.2.1 ログから見えない情報	66
5.2.2 監視にならない測り方の導入	67

5.2.3 プライバシー保護の観点	68
5.2.4 分析・データ貯蔵コスト	69
5.3 展望	69
ヒアリング先自治体.....	70
注.....	71
用語集.....	72

0. はじめに

本章では、本調査の背景、目的、問題意識を整理し、行政 DX¹における効果検証がなぜ重要となっているのかを明確化する。行政環境の変化や DX 推進に伴う新しい行政運営の要請を踏まえ、従来手法では捉えきれなかった課題を提示することで、本研究全体の論点と意義を示す。あわせて、人口減少社会・財政制約・技術進展といったマクロ環境を踏まえ、行政 DX の成果をどのように評価すべきかを問題提起として位置づける。

0.1 調査の目的

本調査は、行政 DX の推進に伴い、施策の実効性を客観的かつ継続的に評価するための成果指標を整理するとともに、それらを低負荷・継続的・比較可能な形で測定するための手法と運用条件を検討することを目的とする。特に、時間、件数、差戻し回数、コスト等の各種数値を可視化する手法を整理するとともに、従来の実測・申告・ヒアリング・アンケート等の有効性を踏まえつつ、測定を追加作業として外付けせず、業務の中で自然に生成されるログや申請データ等を活用する測定設計と運用条件を検討することを主眼とする。

あわせて、職員の追加的な記録負担を抑えながら、継続的な調査・測定を可能とする観測・運用条件を検討し、自治体横断での比較・活用を可能とするデータ活用スキームの方向性を整理する。最終的には、複数自治体で共通して用いることのできる「指標」や「測定方法」を示すことで、政策立案や業務改善に資する信頼性の高いエビデンス²を提供し、自治体業務の効率化を促進することを目指す。

0.2 背景：行政 DX の現状と課題

国においては、デジタル庁の創設以降、標準化・クラウド化を柱とするデジタル社会の実現に向けた政策が推進され、ガバメントクラウドの構築、行政システムの共通化、デー

¹ 行政 DX は多岐にわたる用語であるが、ここでは行政分野において、デジタル技術を活用し、行政サービスや業務プロセスのあり方を見直すことで、住民・事業者の利便性向上、職員の負担軽減、組織運営の改善を図る取組みを指す。

² 本報告書において「エビデンス」とは、手続き時間や業務プロセス、施策の実態を客観的に把握するためのファクト（広義のエビデンス）を指す。ただし、本調査研究が提唱するデジタル測定の枠組みは、施策の反実仮想分析や自治体間比較など、狭義のエビデンスを生成するための土台ともなりうるものである。なお、狭義の意味で用いる場合は、本文中でその旨を明記する。

タ連携基盤の整備が進められてきた。一方で、多くの自治体は財政制約や人口減少といった厳しい環境下にあり、住民の生活様式や価値観の多様化も相まって、行政サービスには迅速性・利便性がこれまで以上に求められている。こうした背景のもと、行政 DX の推進は不可欠な政策課題となっている。加えて、クラウド、API、生成 AI 等の技術進展により行政業務のデジタル化が急速に進み、業務プロセスや成果がデジタルデータとして蓄積される環境が整いつつある。

しかし、DX の進展に比して、その成果の可視化・定量化は十分に進んでいない。現状では、オンライン化率、マイナンバーカード普及率、テレワーク導入率など、あくまで施策の「実施状況」を示すアウトプット指標が中心であり、それらが行政サービスの質や住民利便、内部業務の効率化にどのような成果をもたらしたのかは体系的に示されていない。こうした課題は 2010 年代から指摘されてきたものの、自治体間で統一的な評価指標が確立されているとは言い難く、多くの自治体では導入件数や稼働状況といった実施状況指標に依存したままである。すなわち、行政 DX の実効性を測定するための成果指標や評価手法が未整備であることが、施策評価の根本的な課題となっている。

もっとも、行政 DX の効果検証に参照可能な先行知見が存在しないわけではない。EBPM や行政評価の文脈では、ロジックモデル、アウトプット・アウトカム指標、リサーチデザイン等を通じて、政策効果の発現経路や評価枠組みを整理する知見が蓄積されてきた。また、デジタル公共サービス評価や公共部門測定論においても、サービス品質、利用者満足度、政府活動のインプット・プロセス・アウトプット・アウトカムの区分など、成果をどのように捉え、解釈するかに関する考え方が示されている。

しかし、これらの知見は、評価枠組みや指標、分析手法を整理するうえで有用である一方、それらを実際に機能させるためのデータ取得・蓄積の設計については、なお検討の余地がある。特に、評価に必要なデータを低負荷かつ比較可能な形で取得・蓄積するための実務上の設計論は、十分に体系化されているとは言い難い。本研究は、この不足点に着目し、行政 DX の効果検証を、単発の実測や事後的な説明ではなく、行政実務の中で継続的に把握可能な測定として捉え直す。

0.3 行政 DX の効果検証が困難な理由

行政分野における DX の効果検証が困難である理由としては、主に次の五点が考えられる。一般に、効果検証を適切に実施するためには、妥当な指標の設定と、それに基づく測定設計が不可欠である。しかし行政 DX においては、指標の定義や標準仕様が十分に整備されていないことに加え、それらを日常業務の中で継続的に測定し、比較可能なデータとして蓄積する仕組みが十分に設計されていない。その結果として、施策導入前後の比較や、自治体間・組織間の比較が難しくなり、厳密な効果検証を妨げている。さらに、行政

DX は対象とするサービスや業務の性質が多様であり、単一の基準で一律に評価することが難しいという構造的課題も有している。

なお、本報告では、行政 DX の取組によって生じた変化を把握し、その妥当性を確認したうえで業務改善や政策立案に活用する一連の営みを「効果検証」と総称する。「効果測定」「効果評価」等の類語との使い分けについては、巻末の注 1 に整理する。

0.3.1 指標の未整備

第一に、指標の未整備が挙げられる。これは、効果検証に用いる指標について、統一的な定義や標準仕様が確立していないことに起因する。たとえば、「処理時間の短縮」「利用者満足度の向上」「業務負担の軽減」といった効果はしばしば行政 DX の成果として言及されるが、それらをどのように定義し、どの単位で測定するのかが組織ごとに異なる場合が多い。

また、成果指標（アウトカム）と活動指標（アウトプット）が混在しやすく、何ををもって「効果」とみなすかについて共通理解を形成しにくい。そのため、評価枠組みそのものが安定せず、継続的・横断的な効果検証が困難となる。

0.3.2 データ不足

第二に、指標に関するデータが十分に蓄積されていないことが挙げられる。行政現場では、従来の紙ベース業務や個別最適化された情報システムが長く運用されてきた結果、効果検証に必要なデータがそもそも記録されていない場合が少なくない。たとえば、申請件数は把握されていても、処理に要した実作業時間、差戻し回数、問い合わせ件数、利用者の離脱率などの詳細データは取得されていないことがある。

さらに、データが存在する場合でも、形式が統一されていなかったり、入力ルールが部局ごとに異なっていたりするため、比較可能なデータとして活用しにくい。このようなデータ不足は、効果検証の前提条件そのものを弱める要因となる。

0.3.3 前後比較が困難

第三に、データが存在した場合であっても、施策導入前後で比較可能な形のデータが存在しないことが挙げられる。行政 DX は、単に手続を電子化するだけでなく、業務フローの見直し、制度運用の変更、組織体制の再編などを同時に伴うことが多い。そのため、導入前と導入後では、業務の内容や処理対象そのものが変化してしまい、単純な前後比較が成立しにくい。

また、制度改正や社会情勢の変化、申請件数の増減など、DX 以外の外部要因も結果に

大きな影響を与える。したがって、仮に導入後に処理期間が短縮したとしても、それが DX による効果なのか、他の要因によるものなのかを切り分けることは容易ではない。こうした事情から、行政 DX の因果的効果を厳密に把握することは難しい。

0.3.4 自治体間比較が困難

第四に、自治体間や組織間での比較が困難であることが挙げられる。行政サービスは、人口規模、年齢構成、産業構造、地理的条件、財政状況、職員数、既存システムの構成など、多様な条件の下で提供されている。そのため、同じ手続を対象とした DX であっても、期待される効果や導入コスト、実際の運用成果は大きく異なりうる。

たとえば、大都市自治体と小規模自治体では、利用件数や問い合わせ対応の性質が異なり、同一の指標で単純比較すると実態を見誤る可能性がある。さらに、先進的な自治体ほど DX に積極的に取り組む傾向があるため、比較対象の選定自体にバイアスが生じやすい。したがって、自治体間比較³を通じて一般化可能な知見を得るには、前提条件の差異を十分に考慮した慎重な設計が必要となる。

0.3.5 行政 DX の対象領域の異質性（G2C／G2B／G2E／G2G）

第五に、行政 DX の対象は、住民向けサービス（G2C: Government to Citizen）、事業者向け手続（G2B: Government to Business）、職員向け内部業務（G2E: Government to Employee）、自治体間連携（G2G: Government to Government）といった異なる領域にまたがっており、それぞれで目的、利用者、業務構造が大きく異なることが挙げられる。

これらの領域では、重視される成果も異なる。たとえば、住民向けサービスでは利便性やアクセス性が重要となる一方で、事業者向け手続では専門性や法令適合性、職員向け業務では内部統制や業務品質、自治体間連携では情報連携の正確性や迅速性が重視される。このように、行政 DX は対象ごとに評価軸が異なるため、単一の評価指標で横断的に測定することが難しい。

G2C：住民向け手続の多様性

G2C 領域では、住民が利用する手続の種類が多岐にわたり、その頻度や緊急性、必要書類、本人確認の方法なども手続ごとに大きく異なる。たとえば、住民票の写しの交付申請のような定型的で頻度の高い手続と、子育て・介護・福祉関連のように個別事情への配慮

³ 本報告書において自治体間比較は、自治体間の優劣を評価することではなく、互いの改善や学習を促進するための比較分析を意味する。

が求められる手続とでは、DX による効果の現れ方が異なる。

また、住民の年齢、デジタルリテラシー、利用端末の環境などの違いも、利用率や満足度に影響を与える。そのため、G2C の効果検証においては、単にオンライン利用率だけを見るのでは不十分であり、手続の性質や利用者属性を踏まえた多面的な評価が求められる。

G2B：事業者向け申請の専門性

G2B 領域では、許認可、届出、補助金申請、入札・契約関連手続など、専門性の高い申請が多く含まれる。これらの手続では、処理速度だけでなく、記載内容の正確性、法令・制度要件への適合性、審査の一貫性などが重要な評価要素となる。

さらに、事業者の規模や業種によって必要とされる対応が異なり、同じ電子申請システムであっても利便性の評価は一律ではない。したがって、G2B における DX の効果を測る際には、受付件数や処理日数の短縮といった定量指標に加え、差戻し率、問い合わせ件数、入力負担、再提出の有無など、専門的業務の実態に即した指標を組み合わせる必要がある。

G2E：職員業務の内部プロセスの複雑性

G2E 領域では、行政職員が行う内部業務の効率化や高度化が主な対象となるが、その効果検証は特に難しい。内部業務には、文書管理、決裁、照会回答、予算執行、人事、庶務など多様なプロセスが含まれ、それぞれが複数部門にまたがって連携していることが多い。そのため、ある一部の業務が電子化されたとしても、全体最適としてどの程度の効果があったのかを定量的に把握しにくい。

また、職員業務の改善効果は、単なる作業時間の短縮だけでなく、ミスの削減、引継ぎの容易化、意思決定の迅速化、ナレッジ共有の促進など、質的な側面として現れる場合が多い。こうした効果は数値化が難しく、短期的な指標だけでは十分に捉えられない。

G2G：自治体間連携と標準化依存性

G2G 領域では、国・地方自治体・外郭団体など複数の行政機関の間で情報連携や事務連携を行うことが中心となる。この領域では、個別組織内の改善だけでなく、連携先とのデータ形式の統一、システム仕様の整合、権限管理、法制度上の整理などが不可欠である。

そのため、効果は単独組織の努力だけでは発現しにくく、相互運用性や標準化の達成度に大きく左右される。また、G2G の効果は住民から直接見えにくく、内部的な業務効率化や照会負担の削減として現れることが多いため、成果を可視化しにくいという課題もある。

0.4 調査スコープ

本調査は、行政 DX の成果を適切に把握するための指標体系および測定手法を整理したうえで、最終的には、ログ等のデジタル測定⁴による「測定負荷の少ない効果検証」の実現可能性を明らかにすることを目的とする。こうした目的を踏まえ、本調査のスコープは以下の三点を中心に構成する。

第一に、行政 DX による成果をいかなる観点から把握すべきかについて整理する。具体的には、住民、事業者、職員、行政組織といった多様なステークホルダーに生じるアウトカムを中心に、指標体系を整理する。ここで対象とする指標には、業務時間、コスト、満足度、利便性向上といった概念的指標に加え、利用状況、離脱率、処理時間、経費の前後比較など、ログデータを含む具体的指標も含める。

第二に、これらの指標をどのように測定するかという「測り方」の整理を行う。従来型の実測、アンケート、聞き取り調査等のアナログ手法に加え、ログデータ分析やプロセスマイニングなど、デジタルデータを活用した自動的かつ継続的な測定方法を対象範囲とする。とりわけ、第3章で扱う「測定負荷の少ない効果検証」の観点からは、ログが有する全件性、連続性、自動取得可能性といった特性が、既存の測定手法の制約をいかに克服し得るかを検討する。

第三に、本調査では、既存の評価フレームワークを参照しつつ、行政 DX の特性に適合した新たな評価構造を検討する。特に、第4章で提案する「ログを活用した効果検証モデル」および「業務ログに基づく処理時間・滞留状況の把握モデル」を中心に、行政価値を継続的かつ自動的に可視化する継続観測を支える枠組みの整理を射程に含める。

なお、本調査は、ログ等のデジタルデータを活用した低負荷で継続的に把握する仕組みの構築可能性を主たる検討対象とするが、これをもって行政 DX の成果をログのみで把握し得るとみなすものではない。行政 DX の成果のうち、処理時間、件数、差戻し回数等についてはログによる把握が有効である一方、システム外作業、例外対応、制度運用上の判断、住民・職員の主観的评价等については、実測、ヒアリング、アンケート等による補完が必要となる。このため、本調査では、各手法の役割分担を前提としつつ、継続観測と比較可能性を支える土台としてログを位置づける。また、ログの活用は、個々の職員の勤務状況や能力を評価・監督することを目的とするものではなく、業務プロセス上の課題を把握し、職員の負担軽減及び業務改善につなげるために行うものである。

⁴ デジタル技術によって生成・記録されるデータを活用し、業務やサービスの実行状況を継続的に把握すること。

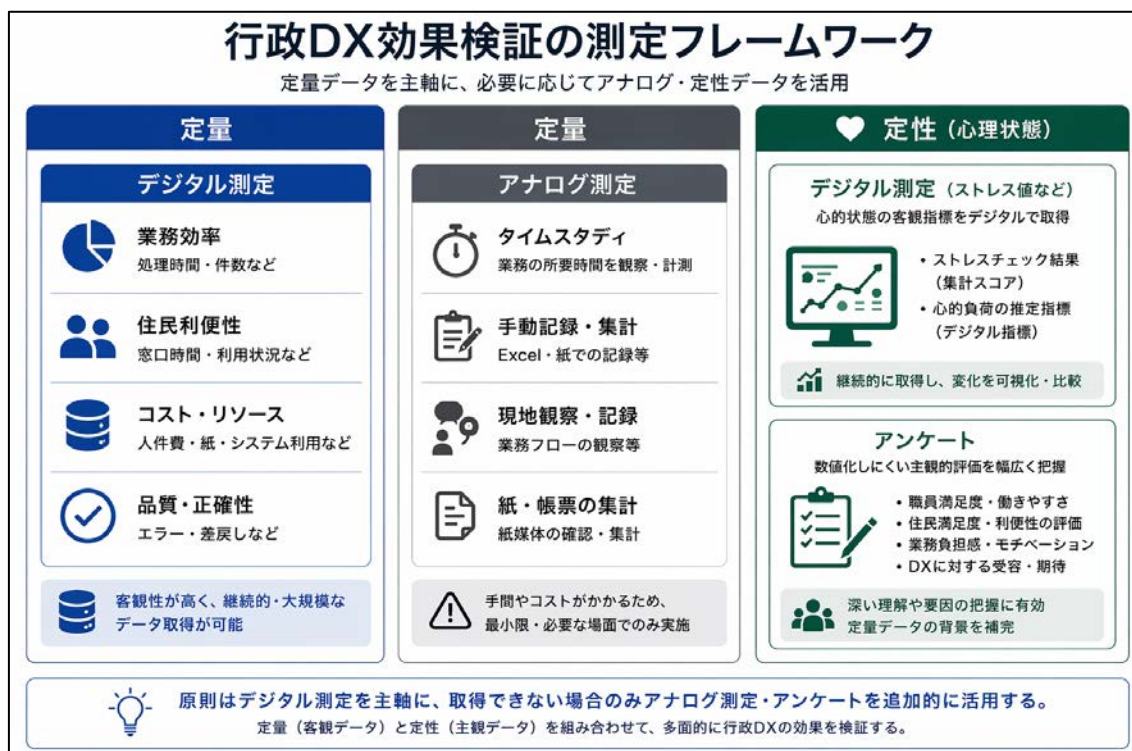


図 1 本調査研究の提唱する概念図

1.行政 DX における効果検証の枠組み

行政 DX の成果を適切に把握するためには、「何を測るのか（指標）」「どのように測るのか（測り方）」「得られた結果をどのような枠組みで価値として整理するのか（効果検証）」の三つのレベルを明確に区別して整理する必要がある。本章では、この三層構造にもとづき、行政 DX の効果検証に関わる要素を体系的に検討する。

1.1 指標の概要

指標とは、行政 DX の取組によって生じる成果をどの側面から把握するのかを定義するものである。従来は、導入件数や稼働状況といったアウトプット指標が中心であったが、行政 DX がもたらす価値は、住民利便性、職員負担、組織運営の改善など多岐にわたる。そのため、住民・事業者・職員・行政組織といった多様なステークホルダーに生じる変化を捉えるアウトカム指標が近年とくに重要視されている。施策評価においては、こうした成果を体系的に整理し、施策目的に応じて適切な指標を設定することが不可欠となる。本節では、行政 DX の成果を把握するうえで代表的な指標区分を概観し、それぞれの特徴を整理する。

1.1.1 指標の定義と範囲

行政施策の評価では、一般にインプット・プロセス・アウトプット・アウトカムの四層構造が用いられる。

- インプット：投入される資源（人員、予算、システム等）
- プロセス：業務手順や作業工程
- アウトプット：提供されたサービス量や対応件数
- アウトカム：住民・職員・組織の変化として現れる成果

行政 DX の効果検証は、従来アウトプットに偏重する傾向があったが、DX の進展に伴い、利便性や業務負担の軽減といったアウトカム指標の重要性が高まっている。

1.1.2 指標の分類

行政 DX の成果は多面的であるため、評価指標についても複数の観点から整理する必要がある。本節では、行政 DX の効果検証で代表的に用いられる 5 つの指標区分について概観する。

時間指標

時間指標は、行政 DX における最も一般的で比較可能性の高い指標である。処理時間、工程別所要時間、待ち時間、年間業務量、業務期間などが該当し、業務効率化の程度を直接的に測定できる。DX によるプロセス改善の成果が明確に表れやすく、多くの自治体事例で中心的な指標として採用されている。

費用指標

費用指標は、業務に要する人件費、外部委託費、運用コストなどを算定するものであり、財政的効果を把握するために用いられる。特に行政 DX では、時間削減量に人件費単価を掛け合わせる形でコスト削減を推計するケースが多い。短期的な削減額に限らず、職員配置の最適化や歳出構造の改善といった中期的観点も重要である。また、業務時間の短縮を人件費単価によって換算した値は、直ちに人件費又は予算の削減額を意味するものではなく、削減された業務時間を金銭単位に換算し、コストの観点から把握するための一つの方法である。

利便性・満足度指標（G2C／G2B）

利便性・満足度指標は、住民や事業者が感じる使いやすさ、心理的負担の軽減、サービス品質の改善など、体験価値を捉える指標である。オンライン化、来庁回数削減、案内のわかりやすさなどが影響し、アンケート等の主観的データによって把握されることが多い。時間や費用といった定量指標では捉えにくい価値を補完する役割を持つ。

職員負担指標（G2E）

職員負担指標は、職員の作業量、作業回数、従事割合、手戻りの発生、精神的負担などを捉えるものである。行政 DX は職員業務の省力化を主要目的の一つとしており、業務負担軽減の程度は重要な成果指標となる。直接測定が難しい場合、担当者ヒアリングや自己申告に基づいて推計されることが多い。

組織運営指標（G2G）

組織運営指標は、庁内調整の迅速化、工程数の削減、データ共有の高度化、意思決定の改善など、組織全体の運営効率を捉える指標である。地理空間情報の統合やデータ連携基盤の整備など、複数部署をまたぐ DX 施策において特に効果が現れやすい。

1.1.3 ステークホルダー別指標

行政 DX は多様なステークホルダーに影響を及ぼすため、指標をステークホルダー別に整理することも有効である。特に、G2C、G2B、G2E、G2G の 4 分類で整理することで、施策の対象と効果の範囲を明確にできる。

住民（G2C）

住民の利便性、待ち時間の減少、手続の簡素化、満足度などを測定する。オンライン化や窓口改革の指標として重要である。

事業者（G2B）

申請手続の合理化、来庁負担の削減、申請漏れ・ミスの削減など、事業者向けサービスの改善を捉える。

職員（G2E）

職員の業務負担、作業時間、手戻り、精神的負担などを評価する。

組織・自治体間（G2G）

庁内の協働、データ共有の円滑化、部署間連携、意思決定の迅速化、自治体間比較（ベンチマーキング）などを対象とする。

1.2 測定方法の概要

行政 DX の効果検証では、対象となる業務や導入技術の性質に応じて、さまざまな測定方法が用いられる。測定方法には、実際の作業を直接計測する方式から、担当者への聞き取りやアンケートを通じて推定値を得る方式、さらにはシステム上のログデータを活用する方式まで幅広い手法が存在する。本節では、代表的な測定方法の特徴と、各手法の長所および限界を概観する。もっとも、これらの測定方法は相互に代替するものではなく、対象とする成果やデータの取得可能性に応じて組み合わせて設計する必要がある。処理時間、件数、差戻し回数、工程別待機時間等のようにログ等で観測可能な領域については、業務の中で自然に生成されるデータを用いて低負荷かつ継続的に把握する。一方で、住民の納得感、職員の心理的負担、例外対応の背景、制度運用上の判断など、ログのみでは把握しにくい領域については、実測、ヒアリング、アンケート等の補完的手法を用いて把握する。このように、行政 DX の効果検証では、ログによる継続観測を土台としつつ、定量・定性の補完手法を組み合わせる測定設計が重要となる。

1.2.1 実測

実測とは、ある業務の開始から終了までの所要時間や作業内容を、ストップウォッチ等を用いて直接計測する方法である。長所としては、現場で実際に行われている業務を精緻に把握できる点や、ボトルネックや無駄な工程を具体的に認識しやすい点が挙げられる。

一方で、限界としては、測定に手間と時間を要すること、現場職員に追加的な負担を与えること、また長期間の継続的な測定が困難であることが指摘される。さらに、実測では、観測者の存在そのものが職員や住民の行動に影響を与える観測バイアスが生じ得る。

1.2.2 試算

試算は、限られた観測データやヒアリング情報から、工程別の所要時間・件数・従事割合を組み合わせて業務量を推定する方法である。導入前後のプロセス差分をフローベースでモデル化し、単価や件数を掛け合わせることで時間・費用効果を算出する。前提条件を併記することが実務上の要諦となる。

1.2.3 ログデータ分析

ログ分析は、業務システム等に記録されたイベントから観測可能な範囲で開始・終了・待ちを復元し、所要時間や手戻りの傾向を把握する方法である。ログが十分に取得できる場合には継続観測を設計しやすい一方、ログ欠落・例外処理・記録粒度の差異には補正が伴う。他手法との併用により、測定バイアスの点検が行いやすくなる。

1.2.4 ワークサンプリング法

ワークサンプリング法とは、一定の期間内でランダムにサンプリングを行い、その時点で従事していた業務内容を記録することで、各業務に要している割合を推定する手法である。サンプリング期間中に実施した観測のうち、特定の業務が記録された回数の比率を算出することで、その業務が全体の作業時間に占める割合を推定することができる。

1.2.5 アンケート

アンケートは、住民・事業者・職員の主観評価を把握する手法である。非市場財の価値を捉えるのに有効だが、回答バイアスや設問設計の影響が大きい。

1.2.6 自己申告法

自己申告法は、担当者の実務知見にもとづき作業時間や従事割合を収集する方法である。自己申告によるため、過少・過大申告が生じやすく、継続モニタリングには不向きである。

1.3 指標×測定法による事例の整理

本節では、指標区分および測定方法を、実際の自治体における行政 DX の取組事例に即して整理する。行政 DX の効果検証は、対象となる業務や導入技術に応じて、時間削減、コスト削減、満足度、業務品質など多様な指標が用いられている。また、それらを算定する測定方法についても、実測、ログ分析、シミュレーション、試算、アンケート調査など、目的に応じて使い分けが行われている。RPA、地理空間情報、窓口改革、AI 活用の事例を特定の手法に偏らずに取り上げ、それぞれが採用する指標×測り方の組合せと、適用条件・制約を整理する。これにより、行政 DX の効果検証における実務的なアプローチと、そのバリエーションを具体的に把握することを目的とする。

1.3.1 RPA（Robotic Process Automation）

RPA（Robotic Process Automation）とは、定型かつ反復的な事務作業をソフトウェア上のロボットで自動化する技術である。行政業務では、入力作業や照合作業などの負荷軽減に寄与し、職員の作業時間削減やミス防止に効果を発揮している。本節では、RPA を活用した自治体の具体的な取組事例を整理する。

足立区:住民異動届への AI-OCR・RPA 導入（時間,コスト,ミス×試算）

足立区では、住民異動届の処理において、職員が書類の確認や転記・登録を手作業で行っていたため、業務負担の大きさと処理時間の長さが課題となっていた。この課題に対し、区は区民サービス向上と重点施策への注力を目的として、AI-OCR と RPA を導入し、書類の読み取りとシステムへの登録作業を自動化した。効果検証では、処理時間と人件費換算によるコストを指標として設定し、試算によって導入前後の時間削減量を算定した。その後、削減時間に人件費単価を掛け合わせることでコスト効果を評価した。

泉大津市: AI・RPA・アウトソーシング（時間×試算 → 人件費）

泉大津市では、人口減少と職員数の不足により、窓口対応や入力作業など多岐にわたる事務が職員の大きな負担となり、コア業務に十分な時間を割けない状況が課題となっていた。この課題に対し、業務の棚卸しを実施した上で、AI・RPA・アウトソーシングを組み合わせた業務改革に取り組んだ。

具体的には、相談や面談内容の文書化を AI で自動化する取り組みや、非定型帳票のデジタル化を進め、手作業による入力負荷を削減した。また、統計資料作成や入力チェックなど定型業務については RPA により自動化を図り、郵送や入力などの単純業務は外部委託を活用した。

効果検証では、担当課へのヒアリングをもとに、工程別の処理時間や作業割合を収集し、

人件費と業務量の関係から削減効果を試算する方式を採用した。試算の結果、入力業務だけで年間約 72,000 時間の業務量が確認され、AI・RPA・アウトソーシングの組み合わせにより、年間数万時間規模の業務削減効果が見込まれた。

塩尻市:保育業務の RPA（時間 × ヒアリング+試算）／（コスト × 試算）

塩尻市では、保育園受付窓口業務において、申込書確認やデータ入力、利用調整など多くの工程が手作業で行われていたため、業務負担の大きさと処理期間の長期化が課題となっていた。この課題に対し、電子申請、AI、RPA、外部委託を組み合わせた業務改革を実施し、データ入力の自動化や AI による利用調整、通知書封入業務の外部委託化を進めた。効果検証では、実測ではなく、担当者ヒアリングに基づき処理時間・件数・従事割合等を収集し、工程別業務量を推計する試算方式を用いた。その結果、処理期間は約 1 か月短縮され、年間約 1,200 時間の削減効果の見込みを得た。

1.3.2 地理空間情報の活用

地理空間情報の活用とは、土地・建物・道路などの位置情報をデジタル化し、GIS や 3D 都市モデル等によって可視化・分析する技術を指す。行政業務においては、各種審査や相談対応の効率化、庁内での情報共有の高度化に寄与している。本節では、地理空間情報を活用した自治体の取組事例を取り上げ、その効果を整理する。

板橋区:統合型 GIS の導入（時間×フロー試算）

板橋区は、庁内の地理情報の参照・更新・共有を一元化する統合型 GIS の導入計画を整理している。効果の把握は導入前後の業務フロー比較に基づく想定で、実測ではない。代表例として、情報閲覧 31 → 11 分／建築物名称変更 35 → 15 分／集積所登録 58 → 34 分／投票区見直し 7 時間 → 1 時間の短縮が見込まれる。

1.3.3 窓口業務改善

窓口業務改善とは、住民が行う申請・届出・相談といった対面業務において、手続の効率化や利便性向上を図るためにデジタル技術を導入し、業務プロセスの見直しやサービス提供方法を改善する取組を指す。行政 DX の進展により、申請書自動作成、案内システム、待ち時間短縮など、住民体験の改善と職員負担の軽減を同時に実現できる領域として注目されている。本節では、窓口業務におけるデジタル技術活用の効果を、自治体の具体的な事例をもとに整理する。

橋本市:窓口支援の導入（時間×シミュレーション実測）

橋本市では、窓口業務において申請書確認やシステム入力、各種審査など多くの工程が手作業で行われていたことから、業務負担の大きさと時間外対応の増加が課題となっていた。この課題に対し、RPA と AI OCR を組み合わせた業務改革に着手し、定型入力作業の自動化や申告書類の電子化に向けた読み取り精度の検証、さらに自治体クラウドを活用した他自治体との共同利用検討を進めた。

効果検証では、実測値に加え、担当課ヒアリングによって処理件数・作業時間・プロセス構造を把握し、業務ごとの自動化効果を推計する方式を採用した。その結果、RPA を適用した複数業務で年間合計 約 300 時間以上の削減効果が見込まれ、特に軽自動車税申告書入力などでは単独で 80 時間超の削減を確認した。一方、AI OCR については手書き文字の認識精度に課題が残り、現時点での実務利用は困難と判断された。

つくば市:市民窓口業務改革（時間×実測、満足度×アンケート）

つくば市は、本人確認書類から申請書を自動作成するサービスを導入した。申請書作成に要する時間をストップウォッチで実測し、導入前後の数値を比較することで効果を測定した。その結果、全体として 3 分以上の時間削減効果が確認されたと報告されている。また、アンケート調査を実施することで、使用感や利便性に関する定性的な評価も行っている。

船橋市（時間×実測,手戻り発生数×試算,満足度×アンケート）

船橋市では、住民異動や証明書発行などの窓口業務において、申請書の記入負担や案内のバラつき、入力作業の煩雑化など多くの課題があり、職員の業務負担増大と市民の手続時間の長期化が問題となっていた。これらの課題に対し、申請書ナビゲーションシステムの導入、業務プロセスの見直し、証明事務や入力業務の集約化・外部委託化、さらに ICT を活用した遠隔相談などを組み合わせた業務改革を実施した。効果の検証では、実測に加え、担当者ヒアリングや業務フロー分析から工程別の作業量を算定する方式を用いた。その結果、窓口滞在時間の削減、手戻り件数の約 45%削減、証明事務や入力業務の外部委託による年間約 1,600 万円規模の歳出効率化が見込まれ、市民サービス向上と職員負担軽減の両立につながった。

1.3.4 AI 活用

AI 活用とは、機械学習や画像認識などの人工知能技術を行政業務に適用し、判断業務やデータ処理の高度化、業務時間の削減を図る取組を指す。行政分野では、固定資産税評価、文書分類、問い合わせ対応など、多様な業務に AI が導入されつつあり、職員の作業負担軽減と業務品質の向上に寄与している。本節では、AI を活用した自治体の取組例を取り上げ、具体的な効果とその測定方法を整理する。

さいたま市:固定資産税評価業務への AI の導入（時間×試算）

さいたま市は、2020 年から固定資産税評価業務への AI 導入を開始した。前年との差分処理を AI で代替することで、評価業務に要する時間を 586 時間から 63 時間へと大幅に削減できると試算している。

1.4 評価フレームワーク

行政 DX の成果を適切に把握するためには、施策を実施した後に評価指標を設定するのではなく、施策の計画段階において、どのような価値を生み出し、どの指標でどのように測定するのかをあらかじめ整理しておく必要がある。そのための手法として、従来の行政評価や公共マネジメントの分野では、目的・プロセス・成果を体系的に構造化するための各種フレームワークが用いられている。本節では、これらの評価フレームワークの構造と特性を整理し、施策立案段階における指標設計や測定・分析手法の選定にどのように活用できるかを示す。

1.4.1 事前評価（ex-ante）と事後評価（ex-post）の位置づけ

行政 DX の効果検証は、本来、事前評価（ex-ante）と事後評価（ex-post）の二段階で構成される。

事前評価では、施策導入前に「どのような価値を生み出すのか」を整理し、目的・指標・測定方法・反事実の想定を明確化する。ロジックモデルや PICO などのフレームワークは、この事前整理を体系化するための設計ツールとして機能する。

一方、事後評価では、実測・ログ分析・試算・因果推論⁵（DID, PSM, RDD など）を用いて、事前に定義した指標に基づき成果を検証する。事前評価での設計と、事後評価での検証を往復させることで、行政 DX の効果検証は再現性と妥当性を確保した形で運用される。

本節では、この事前・事後の二層構造を踏まえつつ、行政 DX に適用可能な代表的な評価フレームワークを整理する。

もっとも、EBPM の実践においては、政策判断に先立って証拠を整え、意思決定に資するという本来の趣旨から離れ、すでに決定された施策を事後的に正当化するための「証拠作

⁵ 因果推論とは、観測された変化が特定の施策や介入によって生じたものかを識別するための分析である。

り」に偏重しうる点に留意が必要である。とりわけ、施策の開始前に基礎となる観測データ（事前データ）が取得されていない場合、効果の検証は導入後の数値のみに依存せざるを得ず、比較の基準を欠いたまま「効果があった」とする結論が先行しやすい。このような事後正当化型の運用は、本来 EBPM が備えるべき意思決定への寄与を形骸化させ、測定や評価が説明責任を果たすための形式的な手続にとどまる懸念を生む。したがって、効果検証は、施策の導入判断に先立つ事前段階から、観測対象・指標・ベースライン⁶を設計しておくことが求められ、事後的な検証はその設計に基づいて行われる必要がある。

1.4.2 ロジックモデル

ロジックモデルは、投入資源から最終的な成果（アウトカム）に至るまでの因果連鎖を可視化し、どの段階をどの指標で測定するのかを体系化する枠組みである。行政 DX においては、システム導入という「手段」が、住民利便性や職員負担軽減といった「目的」にどのようにつながるのかを明らかにし、測定ポイントを整理するために有効である。



図 2 ロジックモデルの例

1.4.3 バランス・スコアカード

バランス・スコアカード（BSC）は、組織の戦略を複数の視点から指標化し、戦略マップとして統合的に運用する枠組みである。行政 DX に適用する際は、民間の「財務」視点を公共資源・効率の視点に読み替え、住民・事業者・職員・組織運営のバランスを取る。

1.4.4 公共価値の三角形（Moore's Triangle）

公共価値の三角形は、DX 施策が公共的価値（Public Value）を生み出すかを、正当性と

⁶ ベースラインとは、施策導入前または比較対象とする時点の状態を定量的に記録したものであり、施策実施後の変化量や効果を測定する際の基準値となる。

支持（Legitimacy & Support）、運用の能力（Operational Capacity）という二つの基盤とともに評価する枠組みである。利便性や効率だけでなく、法的・社会的な正当性と実行可能性を同時に満たすことを重視する。

1.4.5 PICO

PICO は評価対象（Population）、介入（Intervention）、比較（Comparison）、アウトカム（Outcome）を事前定義する設計フレームである。DX 施策の実装前に PICO を合意しておくことで、指標と測り方の後付け化を防ぎ、因果推論や CBA の前提を明確にできる。

1.5 分析方法

本節では、効果検証における評価手法を整理する。評価するにあたり、特に重要となる論点は次の二点である。

1. 削減された業務時間が、本来必要とされるコア業務へ再配分されたかどうか。
2. 観測された変化が DX 施策の効果によるものといえるか（因果性の識別）。

本調査では、まず指標定義と測定方法の整理を目的とするため、因果推論手法（DID、PSM、RDD など）を網羅的に実装することは目指さない。ログデータにより、処理時間や手戻り等を低負荷で継続的に観測する「測定」の土台を整えることを主軸とする。一方で、因果推論手法や費用便益分析（CBA）は、必要な観測条件が整備された場合に、Outcome の変化を施策の効果として識別するための選択肢として位置づける。本節では、それら手法の役割と基本的な考え方を概説し、将来的な指標整備や評価設計の参考材料とする。

DID（Difference in Differences：差の差法）

DID（Difference in Differences）とは、観測することのできない「介入が行われなかった場合の介入群の反事実」を、平行トレンドの仮定によって補完する手法である。平行トレンドの仮定とは、介入が行われなかった場合、介入群と対照群が介入前と同じ傾向で推移したであろうと想定するものである。この仮定が成立することで、対照群の変化を「外的要因による変化」とみなし、その分を差し引くことで介入による純粋な効果を抽出することが可能となる。

DID は比較的シンプルな枠組みでありながら、政策分析や行政分野における実証研究で広く用いられている。一方で、平行トレンドが成立しない場合には推定値にバイアスが生じるため、事前のトレンド分析や追加のロバストネスチェックが重要となる。

PSM (Propensity Score Matching: 傾向スコア法)

PSM (Propensity Score Matching) とは、介入群と対照群の個体について、介入を受ける確率を表す「傾向スコア」が近い者同士をマッチングし、比較を行う手法である。このマッチングにより、介入以外の外的要因を統制し、両群の背景条件を可能な限り揃えたうえで介入効果を推定することが可能となる。

RDD (Regression Discontinuity Design: 不連続回帰デザイン)

RDD (Regression Discontinuity Design) とは、介入の適用基準として設定された閾値の前後に位置する個体の挙動を比較し、その不連続な変化をもとに介入効果を推定する手法である。

閾値付近の個体は、介入の有無以外の属性がほぼ連続的に変化すると仮定されるため、実質的にランダムに割り付けられた状況に近いとみなすことができる。このため、閾値前後で観測される不連続な変化を介入効果として識別でき、準実験的デザインとして因果推定において高い信頼性を有する手法とされている。

一方で、RDD にはいくつかの留意点も存在する。第一に、閾値付近でサンプルが十分に存在しない場合、推定が不安定となる。また、対象者が意図的に閾値を操作できる場合（スコアの調整など）、ランダム性の仮定が崩れ、推定の妥当性が損なわれる可能性がある。さらに、RDD は基本的に「局所的な因果効果」を推定するものであり、閾値付近の個体に限定されるため、母集団全体への一般化には慎重な検討が必要となる。

CBA (Cost Benefit Analysis: 費用便益分析)

CBA (費用便益分析) とは、施策の実施に伴って発生する費用と便益を貨幣価値として比較し、施策の経済的妥当性を評価する手法である。行政施策においては、DX の導入費用や運用コストと、時間削減、利便性向上、エラー削減、人員最適化などの便益を金銭換算し、費用を上回る価値が生じているかを判断するために用いられる。

特に行政 DX では、手続時間の短縮や住民満足度の向上など、直接的な市場価格を持たない便益が多いため、それらをどのように貨幣化するかが重要な論点となる。

こうした場合、便益のうち市場で取引されない「非市場価値」を適切に測定する必要がある。非市場価値には、オンライン手続の利便性、移動不要による心理的負担の軽減、職員の事務作業削減による働きやすさなど、行政 DX が生み出す多様な価値が含まれる。これらは単純な費用削減額では捉えきれず、利用者の主観的価値を金額として測定する枠組みが求められる。

CVM (Contingent Valuation Method : 仮想市場評価法)

CVM (仮想市場評価法) は、非市場価値を貨幣換算するための代表的な手法であり、住民・利用者が特定のサービスに対してどの程度の価値を感じているかを、支払意思額

(Willingness to Pay: WTP) または受取意思額 (Willingness to Accept: WTA) として測定するものである。

CVM では、仮想的な市場状況を提示し、「この行政サービスが提供されるならいくら支払ってもよいか」といった問いかけを行うことで、利用者が感じる価値を金額として推定する。行政 DX においては、オンライン化による利便性向上、窓口待ち時間の削減、移動コストの削減、心理的負担の軽減、手続ミスの減少など、非市場価値が効果の大部分を占めるケースが多い。そのため、CBA を行う際には CVM を併用することで、DX が生み出す便益をより包括的に把握することが可能となる。

一方で、CVM には仮想的な質問形式であることに起因するバイアス（仮想バイアス、戦略的回答バイアス）が存在するため、調査設計には慎重な配慮が求められる。また、住民の理解度や提示方法によって回答が変動し得る点も留意すべき課題である。しかし、非市場価値が大きな割合を占める行政 DX の評価においては、依然として有用性の高い手法である。

TCM (Travel Cost Method: トラベルコスト法)

TCM は、主として公園、博物館、自然環境などの公共財に対する利用者の「利用価値 (use value)」を金銭換算するための手法である。利用者がそのサービスを利用するために実際に支払った移動費用や時間コストを観察し、それらを需要曲線の推定に用いることで、利用者が感じている便益を貨幣価値として評価する。

TCM の基本的な考え方は、あるサービスにアクセスするために人々がどれだけのコストを負担しているかは、そのサービスが持つ価値の表れであるという点にある。すなわち、訪問回数と移動コストの関係から需要関数を推定し、需要曲線の下側に位置する消費者余剰を便益として算出する。これにより市場取引が存在しない公共サービスの価値を、実際の行動データに基づき定量化することが可能となる。行政 DX との関係においては、TCM は「オンライン化によって削減された移動コスト」の便益推計に応用できる。例えば、窓口手続がオンライン化されることで、住民は庁舎までの移動時間・交通費を負担せずに手続を完結できる。このような移動コストの削減は、利用者が実際に支払うはずであった費用に基づいて金額化でき、DX による利便性向上の定量的評価として扱うことができる。

TCM の利点は、実際の行動データに基づいているため、仮想的な回答に依存する CVM と比べてバイアスが小さく、比較的安定した推定が可能である点にある。一方で、訪問目的が複数ある場合の分離が難しいことや、オンライン化による代替行動が多様化すると推定が複雑になるなどの課題も存在する。

SEM (Structural Equation Modeling: 構造方程式モデリング)

SEM は、複数の観測変数から潜在変数（例：住民利便性、職員負担、組織能力など）を

構成し、測定モデル（観測指標と潜在概念の関係）と構造モデル（潜在概念間の因果関係）を同時に推定する手法である。行政 DX では、処理時間や差戻し率といった客観指標に、満足度や使いやすさ等の主観指標を重ねてモデル化することで、体験価値と運用実態の関係を統合的に把握できる。

適用にあたっては、概念枠組みの明確化、信頼性・妥当性の検証、サンプルサイズと識別条件の確保、年度・属性・自治体間での測定不変性の点検が前提となる。推定結果の因果解釈は設計依存であり、前後比較や準実験手法で得られた結果と整合させつつ補完的に用いることが望ましい。モデル誤特定のリスクや質問票設計の質に結果が左右されやすい点にも留意する。

SERVQUAL（サービス品質ギャップ法）

SERVQUAL は、サービス品質を「期待」と「知覚」の差として測定する枠組みであり、信頼性、即応性、確実性、共感性、有形性などの次元から総合的に評価する。行政 DX の文脈では、オンライン手続の分かりやすさ、応答の迅速性、本人確認の安心感、個別事情への配慮、画面・窓口の見やすさ等を設問化し、住民・事業者が知覚する品質を定期的に把握することで、優先改善領域を抽出できる。

1.6 現行枠組みの課題

本章までの整理から、現行の評価枠組みには少なくとも三つの構造的な課題が横たわっている。第一に、指標定義・単位・収集方法が統一されておらず、アウトプット（実施状況）とアウトカム（成果）が混在し、自治体間・時系列の比較可能性が損なわれている点である。第二に、評価が導入件数や稼働率といった実施水準に偏重し、住民利便や職員負担軽減などの成果（アウトカム）把握が手薄になっている点である。第三に、実測・ヒアリング中心の測定は現場負担が大きく継続しにくい一方、ログを用いた継続観測には記録設計や補正が不可欠で、運用要件の明確化が求められる点である。さらに、G2C・G2B・G2E・G2G の異質性が横断指標の設計を難しくしている。

1.6.1 指標未整備と不統一

効果検証においては、処理時間、待ち時間、差戻し回数、利用件数、満足度、職員負担など、多様な指標が用いられる。しかし、これらの指標は自治体や事業ごとに定義、単位、集計粒度、収集方法が異なることが多く、同じ名称の指標であっても、実際に測っている対象や範囲が一致しているとは限らない。

たとえば、「処理時間」といっても、申請受付から完了までのリードタイムを指す場合もあれば、職員が実際に作業した時間のみを指す場合もある。また、「利用件数」について

も、申請件数、処理件数、完了件数、アクセス件数などが混在しやすく、比較の前提が揃わない。このように、指標の意味や測定単位が統一されていない場合、自治体間比較や施策導入前後の比較を行っても、その差が実際の効果によるものなのか、定義や集計方法の違いによるものなのかを判別しにくい。

したがって、行政 DX の効果検証では、単に指標を列挙するだけでは不十分であり、各指標について、定義、計算方法、単位、対象範囲、集計期間、データ取得方法を明確化する必要がある。あわせて、指標の定義書やデータ辞書を整備し、年度や担当者が変わっても同じ意味で測定できる状態を確保することが、継続的かつ比較可能な効果検証の前提となる。

1.6.2 アウトカム不足

行政 DX の評価では、オンライン化率、導入件数、稼働率、処理件数といったアウトプット指標が中心になりやすい。これらは施策の実施状況を把握するうえでは有用であるが、それだけでは、DX によって住民、事業者、職員、行政組織にどのような変化が生じたのかを十分に説明できない。

たとえば、手続きがオンライン化されたとしても、住民にとって手続きが分かりやすくなったのか、来庁や問い合わせの負担が減ったのか、職員の作業負担や手戻りが減ったのか、庁内の意思決定や連携が円滑になったのかは、オンライン化率や導入件数だけでは把握できない。また、処理時間や件数の改善が確認された場合でも、それが住民の納得感や職員の心理的負担の軽減といったアウトカムにどのようにつながったのかは、別途確認する必要がある。

したがって、行政 DX の効果検証では、実施状況を示すアウトプット指標に加え、住民利便、事業者負担、職員負担、組織運営、業務品質といったアウトカムを、ステークホルダー別に整理する必要がある。そのうえで、ログ等から把握可能なプロセス指標と、ヒアリング・アンケート等によって補完すべき主観的・質的アウトカムとの関係を明確化し、単なる導入状況の確認にとどまらない評価設計を行うことが求められる。

1.6.3 測定コスト・運用制約

実測・ヒアリング中心の測定は現場負担が大きく継続困難になりやすい。一方、ログ等

の継続観測は記録設計や補正が前提となる。データの取得・利用可能性・負荷・精度・継続性のトレードオフを明示したうえで、目的別の併用設計により運用可能性を高める必要がある。なお、測定コストには、費用や工数だけでなく、観測対象となる職員の心理的負担や受容性の問題も含まれる。ログ等による継続観測が個人の行動監視や能力評価として受け止められないよう、目的限定、個人評価への不転用、最小限収集、匿名化・仮名化、アクセス統制等のガードレールを設計する必要がある。この点は、5.2.2「監視にならない測り方の導入」で詳述する。

1.7 小括：効果検証枠組みの整理

本章では、行政 DX の効果検証を検討するための基本的な枠組みとして、「何を測るか（指標）」「どのように測るか（測定方法）」「得られた結果をどのような枠組みで解釈し、行政価値として位置づけるか（評価フレームワーク）」の三層から体系的に整理した。行政 DX をめぐっては、従来、オンライン化率や導入件数、稼働状況といった施策の実施状況を示すアウトプット指標が中心となりやすく、住民利便性、職員負担の軽減、組織運営の改善といったアウトカムを十分に把握できていないという課題があった。このため、本章ではまず、行政 DX の成果を把握する際に、指標・測定方法・評価枠組みを区別して整理する必要があることを明確にした。

第一に、指標については、時間、費用、利便性・満足度、職員負担、組織運営といった複数の観点から整理を行った。これにより、行政 DX の成果は単なる処理時間の短縮やコスト削減に還元されるものではなく、住民、事業者、職員、行政組織といった異なるステークホルダーに対して、それぞれ異なる価値として現れることを確認した。たとえば、G2C では利便性や待ち時間の短縮、G2B では申請の正確性や再提出負担の軽減、G2E では業務負担や手戻りの削減、G2G では連携の円滑化や意思決定の迅速化が重要な論点となる。このように、行政 DX の効果は対象領域ごとに評価軸が異なるため、単一の指標で一律に測ることは難しく、施策目的や対象業務に応じた指標設計が不可欠である。

第二に、測定方法については、実測、試算、ログ分析、ワークサンプリング、アンケート、自己申告といった代表的手法の特徴と限界を整理した。実測は工程を精緻に把握しやすい反面、観測負担が大きく継続性に乏しい。試算は限られたデータから業務量や費用効果を推定するのに有効であるが、前提条件への依存が大きい。ログ分析は継続観測や工程分析に適している一方、記録粒度やログ欠落、例外処理の存在に左右される。アンケートや自己申告は主観的価値や現場感覚を把握するうえで有用であるが、回答バイアスや過少・過大申告といった課題を伴う。したがって、行政 DX の効果検証では、どの手法にも固有の長所と限界があることを前提に、対象業務、必要精度、現場負担、継続可能性に応じて複数手法を組み合わせる設計が求められる。

第三に、評価フレームワークについては、ロジックモデル、バランス・スコアカード、公共価値の三角形、PICO などを取り上げ、それぞれが施策目的、指標、測定方法、解釈の仕方を接続する役割を担うことを整理した。行政 DX の効果検証は、単に数値を収集するだけでは成立しない。どの施策が、どのような価値を、誰に対して生み出すのかを事前に構造化し、その後に得られた測定結果を位置づける枠組みが必要である。その意味で、評価フレームワークは、指標と測定を行政価値の文脈へ結びつける設計装置として重要である。とりわけ行政 DX では、事前評価（ex-ante）と事後評価（ex-post）を往復させながら、目的、測定、解釈の整合性を確保することが再現性と妥当性を高めるうえで重要である。

以上の整理から、本章で明らかになったのは、行政 DX の効果検証は、単に「数字を取ること」ではなく、「誰にとってのどのような価値を、どの程度の負荷で、どの程度比較可能な形で把握するか」を設計する営みであるという点である。したがって、効果検証は少なくとも三つの条件、すなわち、第一に誰にとっての価値を捉えるのかという目的適合性、第二に現場負担と継続可能性を踏まえた運用可能性、第三に前後比較や横断比較、因果解釈に耐えうる統計的妥当性、によって設計されるべきである。いずれか一つだけを重視しても十分ではなく、これら三条件の均衡の上に、行政 DX にふさわしい効果検証の仕組みを構築する必要がある。

その際、実務上の基本原則となるのは、単一手法への依存を避けた混合アプローチである。短期の現場改善では実測や自己申告の迅速性が有効な場合がある一方、継続的なモニタリングや比較可能性の確保には、ログ等のデジタルデータを活用した継続観測が有効である。また、住民や職員の主観的評価、納得感、使いやすさといった価値については、アンケートやヒアリングによる補完が不可欠である。したがって、行政 DX の効果検証は、データの取得・利用可能性、コスト、観測負荷、意思決定のタイムラインを踏まえながら、複数の指標、複数の測定手法、複数の評価視角を適切に組み合わせるかたちで構想されるべきである。

本章の整理は、特定の測定手法や評価フレームワークを一律に推奨するものではない。むしろ重要なのは、行政 DX の効果検証を設計する際の判断枠組みを明確にし、何をどのような条件の下で測るのかを事前に構造化することである。次章では、この判断枠組みを実務に適用する前提として、自治体業務の構造的特徴と制度的制約を整理し、なぜ行政 DX において効果検証の実装が困難となりやすいのかを検討する。とりわけ、業務の多様性、法制度上の制約、組織横断的な連携、データの不統一、予算や人事運用の問題など、実務上の条件が効果検証の設計と運用にどのような影響を及ぼすのかを明らかにすることが、次章の課題となる。

2.自治体業務と行政 DX の現状

本章は、自治体業務の構造的特徴と行政 DX の実装状況、そして効果検証が困難となる根本要因を明らかにすることを目的とする。自治体業務の分類（コア／非コア、G2C／G2B／G2E／G2G）を踏まえ、業務ごとのデジタル化の進展度合いや制約条件を整理する。そのうえで、標準化・共通化の推進、DX 実装の進展、組織ガバナンス・人員スキルといった内部環境を踏まえ、なぜ行政 DX が進んでも効果検証だけが進展しにくいのか、その構造的背景を提示する。行政評価における事務量の不可視性やアウトカムの把握困難性など、測定を阻む要因を明確にする。

2.1 自治体業務の構造と特徴

自治体は、都道府県・市町村といった複数の種類に分かれ、それぞれが異なる役割と権限を担っている。さらに、各業務は個別の法令・制度に基づく実施義務を負っており、行政は利益を最大化する営利組織ではなく、公共性と公平性を重視する非営利組織である。このため、自治体業務は制度的制約、法的責任、住民サービス提供という複数の要素が重層的に絡み合い、業務構造が本質的に複雑にならざるを得ない。こうした制度的複雑性は、業務プロセスの標準化や可視化を困難にし、結果として行政 DX の効果をどの観点で測定すべきか、またその成果をどの単位で捉えるべきかを判断する際にも大きな影響を及ぼしている。すなわち、自治体の多様性と制度的枠組みの存在そのものが、行政 DX の効果検証を難しくしている要因の一つである。

2.1.1 業務分類（コア／非コア）

行政サービスの担い手が減少する中、限られたリソースを、より住民価値と行政の責任に直結する「コア業務」に重点配分する必要性が高まっている。本節では、コア業務とは何か、どの基準で区分すべきかを明確にする。以下、区分の基本観点を示す。

【区分の基本観点】

- 法定必須性・代替不可能性：法令に基づく実施義務や、首長部局としての権限行使が不可欠か。
- 住民の生命・財産・権利への即時性：不履行による社会的リスクが高いか。
- 公平性・正確性の要求水準：公的責任としての説明可能性・監査可能性が強く求められるか。
- 需要規模・継続性：件数・繁忙度が高く、継続運用が不可欠か。
- 専門性・権限性：行政固有の判断や裁量、資格・検査権限が伴うか。

- 政策目標との直結度：自治体の基本方針（計画・条例等）と直接に成果連動しているか。
- 代替可能性：標準化・民間代替・外部委託・自動化が相対的に容易か。

コア業務／基幹業務／中核業務

行政固有の責任と権限に基づき、住民の権利・安全・公正な執行を担保する中核的業務。代替が難しく、品質保証と監査可能性が重視される。

G2C：戸籍・住民記録、住民基本台帳関連、児童・生活保護等の支給決定、災害時の避難・救援オペレーション、選挙事務の基幹工程。

G2B：各種許認可の審査・決定、入札・契約の適正管理、監督・指導。

G2E：予算・決算、監査・内部統制、審査・裁定、リスク管理（情報セキュリティ・個人情報保護）等。

G2G：国・他自治体との義務的データ連携、統計・届出、広域災害連携、選挙管理委員会・監査委員会事務の中核部分。

非コア業務／周辺業務／補助業務

目的達成を支える間接・補助的機能で、民間代替・外部委託・自動化の余地が大きい業務。品質担保の必要性はあるが、行政固有の権限行使を伴わないものが中心。

G2C：窓口の案内・呼出し・整理券運用、申請書の作成補助、封入・発送、各種周知広報の制作実務。

G2B：説明会運営、定型的問合せ対応、資料配布・記録作成、スキャン・アーカイブ。

G2E：備品・旅費・庶務処理、会議調整、データ転記、帳票作成、研修運営等。

G2G：定型的データ整形・受け渡し、フォーマット変換、重複入力の調整など。

従来、自治体業務の分類は、定型／非定型、対人／非対人といった業務の形態や、処理量（件数・時間）を軸に整理されることが多い。しかし、業務がもたらす負担は、所要時間や処理件数といった量的側面だけでは捉えきれない。たとえば、苦情・クレーム対応や、生活困窮・虐待事案などの精神的緊張を伴う相談業務は、一件あたりの処理時間が短くとも、職員に大きな心理的負担を課す。こうした負担は時間指標には現れにくく、量的効率化の議論からこぼれ落ちやすい。実際、東京都江戸川区児童相談所では、若手職員の休職を未然に防ぐことを目的に、東京都市大学の宮川氏との共同研究として、職員のストレス状態を客観的に把握する取組を進めている。児童福祉司の心拍変動を記録し、自律神経活動を推定し、VR ロールプレイ研修と組み合わせることで、従来のアンケートでは見えにくかった負担を捉えようとしている。測定に対する現場の反応はおおむねポジティブ

であり、結果を個人評価には用いず、管理職が業務配分をさりげなく調整する材料として活用することで、測定への信頼が保たれている。

したがって、業務分類にあたっては、量的負担に加えて「負担の質（心理的・感情的負荷）」という観点を補助軸として加えることが望ましい。この観点は、効果検証の射程を「効率性（時間・件数）」から「職員の負担・健康」へと広げるものであり、ログには痕跡として残りにくい非観測領域（5.2.1）や、ストレス・生理指標の議論とも接続する。

2.1.2 プロセス分類（法令／住民／内部／更新）

本節では、自治体業務を「法令・制度に基づく必須業務」「住民業務（G2C）」「内部審査（G2E）」「データ更新・庁内連携（G2G）」の4類型として整理する。ステークホルダーの観点と対応づけることで、各プロセスの特性と評価軸を明確にし、後続の効果検証設計につなげる。

法令・制度に基づく必須業務（多くが G2G／G2E）

【定義】

個別の法令や制度に基づき、適正手続・公平性・説明可能性が強く要請される基幹的業務である。審査・決定・監査・選挙などを包含し、運用においては正確性と監査可能性を最優先する点に特徴がある。代表例として、各種許認可の審査・決定、入札・契約の適正管理、監査・内部統制、選挙事務の基幹工程などが挙げられる。

住民業務（G2C）

窓口・オンラインを問わず、住民の申請・届出・相談等に直結する対人サービス領域であり、利便性やアクセス性、待ち時間、満足度といった体験価値の管理が重視される。代表的な業務には、各種証明書の交付、住民異動の手続、子育て・福祉関連の申請、税関連の各種届出などが含まれる。

内部審査（G2E）

庁内における文書管理や決裁、各種照会への回答、予算執行、人事・庶務といった職員が遂行する内部プロセスを指す。多部門が連携して進むため、個別工程のみの改善では全体効果が埋没しやすいという構造的特徴がある。主な対象として、回覧・決裁、照会回答、支出負担行為や支払、旅費・備品・庶務処理、文書・契約書管理などが挙げられる。

データ更新・庁内連携（G2G）

国・都道府県・市区町村・外郭団体間の情報連携や、庁内部局間のデータ更新・同期化を担う領域であり、相互運用性や標準化、権限管理への依存度が高い。代表例として、法

定統計・各種届出データの提出、基幹台帳の更新・参照、連携先とのフォーマット整合、庁内 GIS や共有データベースの更新がある。ネットワーク分離（三層ネットワーク等）や権限制御の設計が連携要件に影響する点にも留意が必要である。

2.2 DX 取り組み状況と課題

本節では、自治体におけるデジタル化の到達点と運用上の制約を、標準化・共通化の進展、実装状況の実態、そして人員・スキル・ガバナンスの三つの観点から概観する。そのうえで、東京都の事例を取り上げ、制度設計と運用設計がどのように結びつき、成果の現れ方と評価の測り方にどのような示唆を与えるのかを具体化する。

2.2.1 標準化・共通化の進展

標準化と共通化は、個別最適のばらつきを抑え、横断的な評価や継続的なモニタリングの前提を整える役割を担う。オンライン前提の処理基準や進捗の可視化は品質管理の基礎となる一方で、記録粒度、例外処理、データ連携の要件など、運用設計で解くべき論点は残存する。標準化によって計測可能な領域が広がるほど、指標設計と測り方の更新が不可欠となる。

2.2.2 DX 進行状況

実装は領域ごとに不均一である。住民起点の高頻度手続ではオンライン化が先行し、内部審査や自治体間連携ではガバナンスとデータ品質の要件が高いため、工程の見直しと記録設計を組み込んだ段階的な移行が求められる。デジタル化の効果を十分に発揮するには、単なる手段置換にとどめず、業務フローの再設計、添付書類の省略、モバイル完結性など、体験設計を含む運用最適化が必要となる。

2.2.3 東京都の事例

デジタル庁による全体方針のもとで行政のデジタル化が進む一方、各自治体でも独自の取組が展開されている。東京都は、東京デジタルファースト条例に基づき、行政手続の原則デジタル化を推進している。令和六年度から八年度を対象とする東京デジタルファースト推進計画の第二期では、対象手続の棚卸を四半期ごとに更新し、オンライン化の進捗をダッシュボードで公表する運用が制度化されている。これにより、計画対象や KPI 管理対象の可視化を通じて全庁横断で実装を管理している。

同計画の柱として、サービスの質の向上を「より早く、よりシンプルに、より使いやすい

く」の三原則で定義し、行政手続デジタルクオリティマネジメントによって品質基準を見える化しながら改善サイクルを回す枠組みを導入している。年間申請件数が一万件以上で全体件数の九割超を占める手続を重点手続として指定し、令和八年度までにレベル二の品質水準への到達を目標とし、その達成度を行政手続デジタルクオリティマップとして公表していく方針である。

処理の透明性と迅速性を担保するため、東京都は手続ごとに標準処理期間を設定し公表している。デジタル前提での運用に合わせ、オンラインで手続を行った場合のオンライン標準処理期間を新たに設定し、処理状況の可視化と組み合わせることで、申請者の体感時間の短縮も含めて管理する考え方を示している。

さらに、添付書類の省略や電子化、行政機関間の情報連携を通じて、ワンスオンリーとワンスストップの実現を段階的に進めている。事業者データベースを核とする手続サクサクプロジェクトでは、法人番号や所在地などの法人情報を一度登録すれば複数手続で再利用できる仕組みを整備し、事業者の入力負担と庁内事務の双方の削減を狙う。まずは補助金等の分野から実装を進め、全庁への展開を見据えている。

こうした取り組みの中で、東京都は DX 施策の成果指標として「手取り時間」概念を導入している。「手取り時間」とは、1 日 24 時間のうち住民や職員が自由に使える時間を DX の成果指標として捉え直したものであり、2025 年 1 月の都知事年頭挨拶で初めて提示された。同年 3 月策定の「シン・トセイ X」では、都民・事業者・職員の「手取り時間」を増やすことが戦略目標として位置づけられている。

「2050 東京戦略」では、生成 AI の活用による年間 50 時間の業務時間短縮や、テレワーク週 1 日実施による年間約 79 時間の通勤時間削減といった具体的な試算が提示されており、職員アンケートや総務省統計局「社会生活基本調査」を用いたモデル換算により算出されている。実装面では、2023 年設立の GovTech 東京が中核となり、行政手続のデジタル化やデータ活用基盤の整備が進められている。2026 年 4 月からは都職員約 6 万人を対象とする生成 AI 共通基盤「A1（えいいち）」の本格運用が開始されており、業務の効率化と時間創出を具体的実現するための環境整備が進展している。

2.3 効果検証が困難となる要因

以上、行政業務を俯瞰した結果、前節までで整理したとおり、効果検証が難しいのは、指標・データ・比較可能性・継続運用の前提が揃いにくいからである（0.3 節、1.6 節）。本節では、同じ結論を繰り返すのではなく、その背景を自治体業務の構造と制度・運用条件（業務量の不可視性、PDCA 不全、予算執行・人事ローテーション等）として分解し、なぜ現場で実装が止まりやすいのかを説明する。

2.3.1 業務量の可視化不足

多くの業務で、工程別の処理時間・待ち時間・手戻り発生といった評価に不可欠なデータが体系的に記録されていない。紙起点の運用や個別最適化されたシステムが長く併存してきた結果、記録自体がない、または記録形式・入力ルールが部局で不統一であるため、横並び比較や継続観測が困難となる。

さらに、指標の定義が組織横断で整っていないため、アウトプット（実施状況）とアウトカム（成果）の混在が生じ、何を「効果」とみなすかの共通理解が形成されにくい。これにより、データ収集の範囲・単位・期間が各部署の判断に委ねられ、評価の再現性が損なわれる。

2.3.2 成果の定量化が困難

第二に、行政 DX による成果は多元的であり、単一の指標によって定量化することが難しい。行政 DX の効果は、住民にとっての利便性向上、事業者にとっての申請負担の軽減、職員にとっての業務負担の軽減、組織にとっての意思決定や連携の円滑化など、複数のステークホルダーに異なる形で現れる。そのため、処理時間や件数のように比較的測定しやすい指標だけでは、行政 DX が生み出す価値の全体を十分に捉えることはできない。

また、行政 DX の成果には、数値として把握しやすいものと、定量化が難しいものが混在している。たとえば、申請処理時間、差戻し回数、来庁回数、作業件数などは、システム上の記録や業務ログから把握しやすい。一方で、住民が感じる安心感や分かりやすさ、職員の心理的負担、相談対応における納得感、組織内での知識共有や判断の質といった成果は、ログや件数だけでは十分に把握できない。これらは行政サービスの質に関わる重要な成果であるにもかかわらず、定量指標として表現しにくいいため、評価の対象から抜け落ちやすい。

このように、行政 DX の成果は、効率性、利便性、負担軽減、業務品質、組織運営といった複数の価値から構成されるため、単一の数値に集約することは適切ではない。むしろ、対象業務やステークホルダーに応じて、複数の指標を組み合わせ、ログで把握可能な成果と、ヒアリング、アンケート、実測等によって補完すべき成果とを区別して設計することが求められる。

2.3.3 PDCA 不全（特に G2E・G2G 領域）

内部業務（G2E）は多部門にまたがる工程が前提で、個別工程の改善を観測しても全体最適としての成果に結び付きにくい。自治体間連携（G2G）はデータ形式や仕様の整合性、権限管理、法制度の整理など外部条件への依存が大きく、単独組織の取組だけでは効果が可視化されにくい。この構造により、検証（Check）が弱まり、改善（Act）へつながる合意や措置が滞る。

観測設計の欠落も影響が大きい。導入前の基準値がない、集計粒度が粗い、タイムスタンプの定義が不統一、案件 ID 等の連結キーが欠落して工程横断の追跡ができない、といった状況では前後比較や因果効果の識別が不安定になる。実測・ヒアリング中心の測定は負担が大きく継続しにくい一方、ログを用いた継続観測には記録設計と補正が必要で、要件が曖昧だと単発の事後評価にとどまりやすい。

克服には、年度境界をまたいで連続的に観測する計画（四半期基準で繁忙・閑散期を跨ぐ）の策定、調達段階でのログ項目・粒度・語彙、エクスポート／API、監査ログ、スキーマ版管理、ID 連結、匿名化・仮名化、データ持ち出しの出口条項の要件化が不可欠である。併せて、データ辞書・指標の定義書・メタデータの版管理と観測手順の文書化を進め、「監視にならない測り方」を明文化し、記録から分析までの費目を年次固定費として確保する。これらは第 3 章および第 4 章の運用モデルで具体化する。

2.3.4 予算執行と人事ローテーション

年度単位の予算・調達サイクルは、測定と可視化を単発の事業に閉じ込めやすい。年度外の観測や繁忙・閑散期を跨ぐ連続観測が設計されず、基準値や時系列の連続性が途切れやすい。さらに、調達仕様にログ出力、エクスポート／API、監査ログ、スキーマの版管理、ID 連結、匿名化・仮名化、データ持ち出しの出口条項などを織り込まないと、改修や更改のたびに指標定義や記録粒度が変質し、比較可能性が損なわれる。

人事ローテーションは組織上の合理性がある一方で、指標定義・観測手順・メタデータ管理の知識を断絶させやすい。同名の指標でも意味が年度ごとに揺らぎ、工程横断や自治体間の比較が不安定になる。ベンダー固有形式への依存やデータ持ち出しコストが高い場合、仕組みをまたぐ長期運用はさらに難しくなる。

対策として、年度境界をまたぐ観測計画を前提化し、四半期単位の連続観測と繁忙・閑散期を跨いだデータ取得を基準化する。調達段階で前記の計測要件を明示し、部局横断の合議体でデータ辞書や指標の定義書、メタデータを版管理する。観測手順、欠測補正、異常時対応、可視化仕様を引継ぎ資産として文書化し、記録・品質点検・可視化・分析の費目を年次固定費として位置付けることで、継続観測と改善の循環を安定化できる。

2.4 小括

本章では、行政 DX の効果検証を実装する前提として、自治体業務の構造的特徴、DX の取組状況、そして効果検証が困難となる根本要因を整理した。目的は、単に「効果検証が難しい」と述べるのではなく、どのような業務特性と制度・運用条件が、観測設計と継続運用の前提を揃えにくくしているのかを明らかにし、次章以降のモデル設計で織り込むべき制約条件を特定することにあった。

第一に、自治体業務は、法令・制度に基づく義務的業務を多く含み、公共性・公平性・説明可能性・監査可能性といった要請が業務プロセスに内在しているため、業務構造が本質的に複雑であることを確認した。これに対処するため、本章では業務を「コア／非コア」の観点、および「法令・制度に基づく必須業務」「住民業務（G2C）」「内部審査（G2E）」「データ更新・庁内連携（G2G）」というプロセス類型として整理し、どの領域で何が成果として現れやすいか／何が評価軸になりやすいかを見通せる形にした。こうした整理は、後段で指標・ログ設計を行う際に、対象範囲と観測単位（案件・工程・部署等）を明確化するための基礎となる。

第二に、自治体における DX 実装は領域ごとに不均一であり、住民起点の高頻度手続ではオンライン化が先行する一方、内部審査（G2E）や自治体間連携（G2G）ではガバナンスやデータ品質、権限管理等の要件が高く、工程の見直しと記録設計を組み込んだ段階的な移行が求められる点を整理した。加えて、東京都の事例は、手続の棚卸を四半期ごとに更新し可視化する運用や、品質基準（クオリティマネジメント）と処理期間の公表を組み合わせることで、単なるオンライン化率にとどまらず、体験品質や処理の透明性を管理対象に含めていることを示しており、制度と運用の設計が、測れる対象（観測可能性）そのものを規定するという示唆を与える。

第三に、前節までで整理したとおり、効果検証が難しい要因は「指標・データ・比較可能性・継続運用の前提が揃いにくい」という点に集約される（0.3 節、1.6 節）。本章では同じ結論を繰り返すのではなく、その背景を自治体業務の実務条件として分解し、(1) 工程別の処理時間・待ち時間・手戻りといったデータの記録が体系化されにくいこと、(2) 内部業務や連携業務では工程連鎖・外部条件への依存が大きく、PDCA の「検証（Check）」が弱まりやすいこと、(3) 年度単位の予算・調達サイクルや人事ローテーションによって観測期間と知識が分断され、同名指標でも意味が揺らぎ、長期比較が不安定になること、を中心に整理した。これらは測定の技術論だけで解ける問題ではなく、調達仕様・ログ項目・語彙・メタデータ管理・引継ぎ資産化といった制度・運用設計を前提として、はじめて継続観測と改善循環に接続し得る。

以上より、第 2 章で明らかになったのは、行政 DX の効果検証は「指標を決める」「データを集める」といった個別作業では成立しにくく、業務類型ごとの構造と、年度・調達・人事を含む運用条件を前提化したうえで、観測の単位・粒度・連結キー・継続体制を設計する必要があるという点である。次章以降では、本章で整理した制約条件を前提として、ログ等のデジタルデータを基盤に、低負荷で継続運用可能な測定・分析・改善の仕組みを具体化する。

3. 「測定負担」を増やさないデジタル測定の考え方

本章では、従来の効果検証が抱えてきた負担・制約を整理し、ログデータを中心としたデジタル測定がどのようにそれを克服し得るかを提示する。行政現場では、実測やヒアリングなど、測定自体が負担となる方式に依存してきたため、継続的かつ客観的なデータ収集が難しかった。本章では、ログの「自動性」「全件性」「連続性」という特性が、従来の課題（負担・観測点不足・継続の困難）をどのように解消するのかを示し、プロセスマイニングなどの分析手法が行政評価の精度と再現性を高める可能性を論じる。あわせて、国内外の先行例を踏まえ、デジタル測定の実装可能性を具体的に示す。

3.1 なぜ従来の測定は負担を生み、継続しにくかったのか

前節までで整理したとおり、困難さの論点は既に提示した。しかし本章の目的は、ログを用いた測定の可能性を示す前に、「従来手法ではどの条件が満たせなかったのか」を測定要件として言語化することにある。以下では、負担・観測設計・継続性の三点から整理し、ログがどの課題を構造的に改善し得るかを位置づける。

行政 DX の効果検証は、(1) 測定そのものが現場の新たな負担になる、(2) 比較・検証に必要な観測時点や観測粒度が不足する、(3) 長期にわたる継続観測の設計・運用が困難、という三つの構造的要因に阻まれてきた。これらは単独ではなく相互に関連しており、測定の設計段階で前提条件や制約を明示しない限り、再現性と妥当性を同時に満たすことが難しい。本節では、従来手法が直面してきたボトルネックを三つの観点から整理し、次節以降で論じる「ログを中心としたデジタル測定」による克服可能性を位置づける。

3.1.1 計測が生む現場への負担

効果検証では、測定そのものが目的化し、現場負担を増大させるリスクがある。業務改善を目的として測ろうとしても、実情としては現場への負担を増やすことにつながる可能性がある。

実務の現場では、測定のために別途の記録作業や説明対応を求められる場合、それ自体が新たな業務負担として認識されやすい。改善のために測ろうとしているにもかかわらず、測るための作業が増えることで、かえって本来業務や改善活動に充てる時間が圧迫され、本末転倒となるおそれがある。特に、実測やヒアリングに依存する手法では、観測に協力する担当者側の負担が大きく、継続的な運用が難しくなりやすい。

したがって、効果検証の設計において重要なのは、何を測るかだけではなく、どのよう

な負担構造で測るかである。測定のために新たな業務を付け加える方式には限界があり、可能な限り日常業務の過程で自然に蓄積されるデータを起点として観測を構想する必要がある。

3.1.2 効果検証に耐えうる観測時点の不足

効果を適切に識別するためには、導入前後の観測時点と観測粒度を計画的に確保することが不可欠である。しかし実務では、この設計が事後的かつ断片的になりやすく、因果推論に耐えるだけの観測点が不足する。まず、導入前の基準値が取得されていない、あるいは年度・四半期集計など粗い単位しか残っていないため、前後比較の前提が欠落する場合が多い。

また、導入前の状態が十分に記録されていない場合、改善後の数値だけを示されても、それを何と比較して効果とみなすべきか判断しにくい。その結果、効果検証は定量的な比較というよりも、担当者の感覚や事後的な説明に依存しやすくなり、検証の再現性や説得力が弱まる。

したがって、観測時点の不足は単なるデータ欠如の問題ではなく、前後比較や因果解釈の基盤そのものを欠く問題である。効果検証に耐えうる観測設計とは、導入前後の時点を確認し、必要な粒度で継続的に観測できる状態を、あらかじめ設計しておくことを意味する。

3.1.3 長期的測定の難しさ

効果検証を継続的に運用するには、単発の前後比較とは異なる設計と体制、そして恒常的な資源配分が必要である。実務では、調達・予算・人員運用が年度単位で分断されやすく、予算期間外の観測を計画的に実施することが難しい。その結果、測定が一度きりの実施で終わり、翌年度に継続観測が途切れる。数値が改善サイクルや意思決定に接続せず、検証・改善・再測定の循環が回らないことが、長期的測定の本質的な課題である。

継続が途切れる背景には複合的な要因がある。試行的な事業や単年度の委託の枠内で完結してしまい、記録・保守・可視化・分析といった「計測の運用費」が年次の固定費として措置されないこと。初期の調達仕様にログ出力、エクスポートやAPI、スキーマの版管理、監査ログなど測定に不可欠な条件が内在化されておらず、システム改修や更改のたびに指標定義や記録粒度が変質して時系列の連続性が失われること。人事ローテーションが頻繁な環境で、指標定義・手順・メタデータ管理の知識の引継ぎが不足し、同名指標でも実態的な意味が乖離しやすいこと。ベンダー固有形式への依存やデータ持ち出しのコスト高によって、長期運用が難しくなること。さらに、継続観測が「監視」と受け取られるリ

スクに対して、目的限定・最小限収集・匿名化などのルール設計と合意形成が十分でない場合、運用そのものが停止するおそれがある。最後に、データオーナーやスチュワード、分析責任者といった役割や、定例のモニタリング会議が制度として設計されていないと、指標の更新・品質点検・解釈の共有が行われず、継続の価値が可視化されない。

これらの課題に対処するには、技術的手当てだけでは不十分であり、制度設計と運用設計を一体で見直す必要がある。年度境界をまたぐ観測計画をあらかじめ作成し、四半期単位の連続観測と繁忙期・閑散期を跨いだ観測を基準化すること。調達段階で、ログ項目・粒度・タイムスタンプ定義、エクスポート／API、監査ログ、スキーマ版管理、ID 連結キー、匿名化・仮名化、データ持ち出しの出口条項を要件化すること。部局横断の合議体を設け、データ辞書や指標の定義書、メタデータの版管理を継続運用すること。観測手順、欠測補正、異常時対応、可視化仕様を「引継ぎ資産」として文書化し、人事交代を前提とした持続性を担保すること。目的限定、最小限収集、アクセス統制、監査記録、第三者点検といった「監視にならない測り方」の原則を明文化・周知すること。そして、記録・品質点検・可視化・分析の費目をシステム運用費と同列の年次固定費として位置づけ、PoC 止まりを回避することが求められる。あわせて、自治体横断の共同ベンチマークを推進し、共通語彙と共通指標を活用して定義の変質を抑えつつ、横比較と継続観測の両立を図ることが有効である。

3.2 デジタル測定としてのログデータ

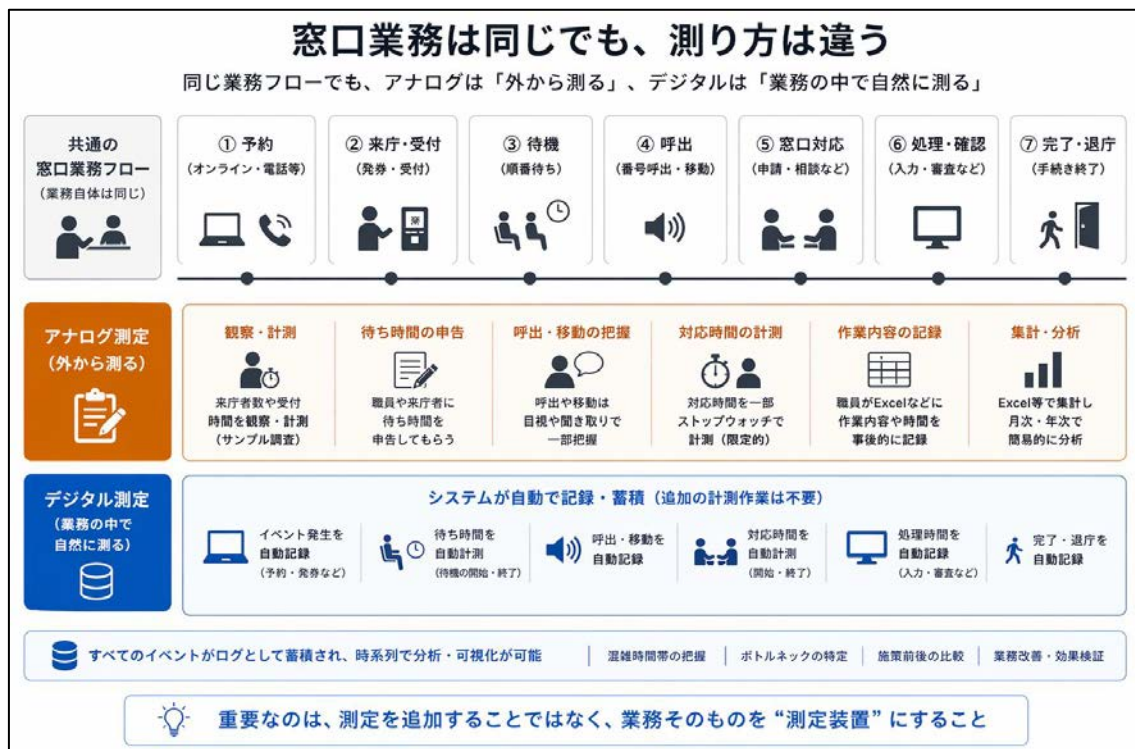


図 3 デジタル測定の考え方

本章で扱うログデータの活用は、個々の職員の勤務状況や能力を評価・監督するためのものではない。業務プロセスにおける処理時間、滞留、手戻り等を把握し、職員の負担軽減及び業務改善につなげることを目的とする

行政 DX の進展に伴い、従来は紙媒体や口頭処理に依存していた行政活動の多くが、業務システム上で遂行されるようになりつつある。その結果、申請の受付、審査の開始、差戻し、承認、完了といった処理工程は、業務の遂行過程においてデジタルデータとして記録されるようになっている。

本研究では、このように行政活動の遂行に伴って副次的に生成されるデータを「プロセス副産物データ」と呼ぶ。そのうち、案件、処理内容、時刻、担当組織等が一定の構造をもって記録され、業務プロセスを時系列に復元し得るものを「ログデータ」と位置づける。たとえば、電子申請の受付記録、ワークフロー上の承認記録、業務システムの操作記録、入退館記録、端末利用記録等は、いずれも業務運用の過程で自然に生成されるプロセス副産物データであり、構造化された時系列記録として取得できる場合には、効果検証に用いるログデータとなり得る。

重要なのは、これらのデータが、特定の測定目的のために職員が別途記録する調査データではなく、業務の遂行そのものに伴って生成される点である。従来の実測やアンケート調査等では、観測対象となる職員による記録作業や回答行為が必要であり、測定そのものが業務負担となる場合があった。これに対し、プロセス副産物データとしてのログデータは、業務の中で自然に生成されるため、追加的な測定行為を抑えながら、行政活動の実行

状況を継続的に把握する基盤となり得る。

3.2.1 ログデータにより把握可能な成果の範囲

ログデータは、行政業務の処理過程に関する履歴情報を継続的に記録することにより、従来の実測やヒアリングでは把握が困難であった業務の実行状況を観測可能とする。ただし、ログによって把握可能な成果は、すべての行政価値を網羅するものではなく、観測可能性の観点から以下の三類型に整理できる。

第一に、ログにより直接的に観測可能な成果として、処理時間、工程別待機時間、差戻し回数、作業件数等が挙げられる。これらはシステム上の処理履歴やタイムスタンプにより全件的に取得可能であり、業務効率性や処理負荷の変化を客観的に把握する指標となり得る。

第二に、ログから代理的に把握可能な成果として、来庁回数の削減や手続完了までの所要日数の短縮等が挙げられる。これらはアウトカムに接近する手がかりとして有用であるが、住民が実際に感じる利便性や安心感、職員の心理的負担の変化といったアウトカムそのものとは区別される必要がある。したがって、代理指標の改善をもってアウトカムが達成されたと結論づけることは適切ではなく、両者の関係はロジックモデル等を通じて仮説として明示されるべきである。

第三に、ログのみでは把握が困難な成果として、住民の納得感、手続の分かりやすさ、職員の心理的負担といった主観的・質的側面が存在する。これらについては、アンケート調査等の補完的手法との併用が必要となる。

したがって、本調査では、ログデータにより観測可能な範囲を効果検証の土台として位置づけるとともに、ログで把握困難な成果については補完的な評価手法の適用を前提とする。

3.2.2 ログが実現する自動計測

ログデータの第一の特性は、行政活動の遂行と同時に自動的に生成されるという「自動計測性」にある。一般に、効果検証におけるデータ収集は、測定対象の行動や作業内容を観察・記録する観測主体の存在を前提としてきた。しかし、このような観測主体を必要とする手法は、測定対象者の業務遂行に影響を及ぼす可能性があるほか、測定対象範囲や観測頻度の制約から継続的な測定が困難となりやすい。

これに対し、デジタル化された行政業務においては、システム操作や処理の履歴が業務遂行の過程で自動的に記録されるため、観測者の介入を必要としない。電子申請や内部決裁の処理過程は、ワークフロー上の操作履歴として時間情報を伴い蓄積され、入退館記録や業務端末の移動履歴も同様に自動的に生成される。このような履歴情報は、測定のための

追加的な記録作業を伴わずに取得可能であり、観測行為そのものが測定対象の行動に影響を与えることを回避し得る。

したがって、ログデータを用いた測定は、従来の実測やヒアリング等の手法に比して、現場への追加的な負担を低減しつつ、業務の実行状況を継続的に観測可能とする特性を有している。

ログデータのように業務遂行の過程で自動的に生成される履歴情報は、個々の職員の能力評価ではなく、業務上の行動や処理プロセスの痕跡として取得される点に特徴がある。このため、処理工程や機能利用状況といった比較的客観性の高い観測対象から段階的に測定を開始することが可能となる。

3.2.3 全件・連続データとしてのログの特性

ログデータの第二の特性は、行政活動の履歴を全件的かつ連続的に記録し得る点にある。従来の効果検証手法では、業務時間の実測やアンケート調査、ヒアリングなど、観測対象の一部を抽出したサンプルデータに基づいて効果の推定を行うことが一般的であった。しかし、このようなサンプリングに依存した測定方法では、対象期間の代表性や観測頻度の制約、繁忙期・閑散期の偏在などにより、統計的に十分な観測点を確保することが難しく、施策導入前後の比較や自治体間の比較を実施する際の制約となり得る。

これに対し、デジタル化された行政業務では、各処理や操作が日常的にシステム上で実行されることに伴い、その履歴が個別案件単位で継続的に蓄積される。このようなログデータは、特定期間のサンプルではなく、業務運用の全体を通じた観測記録として取得されるため、年間を通じた処理件数や案件処理時間、差戻し回数、工程間の待ち時間等を網羅的に把握することが可能となる。

さらに、ログデータは一定の頻度で継続的に生成されることから、施策導入前後の時系列データを長期にわたって蓄積することができる。このような連続的観測は、外的環境の変化や制度改正、需要の季節性といった要因を統計的に統制しながら、施策導入の効果を識別するための前提条件を整えるものとなる。たとえば、繁忙期と閑散期が繰り返される行政手続において、複数年度にわたる処理履歴を取得できる場合には、短期的な断面比較ではなく、導入前後の変化をトレンドとして捉えることが可能となる。

さらに、行政業務の多くが法令や制度に基づく義務的な手続である点も、ログデータを用いた効果検証の適用可能性を高める要因となる。たとえば、住民異動届の受付や各種認定手続、許認可の審査といった行政サービスは、自治体ごとの差異は存在するものの、その基本的な処理の流れは関連法令や通知によって制度的に規定されている。このため、申請、確認、審査、決裁、通知といった業務工程の構造は、自治体間において一定の制度・業務構造の共通性を有していると考えられる。

従来は、こうした制度的に類似した業務であっても、紙媒体を土台とした運用のもとで

は処理過程を工程単位で把握することが困難であり、個別業務の効率性を横断的に比較することは実務上難しかった。しかし、デジタル化された業務環境においては、案件ごとの処理履歴がログとして記録されることで、制度上類似した業務プロセスの実行状況を定量的に把握することが可能となる。このような制度的・業務構造の共通性とログの全件観測性の組み合わせは、自治体間における業務処理効率の比較やベンチマーキングといった分析の実施可能性を高める土台となり得る。

3.2.4 測定行為を伴わないデータ取得がもたらす利点

ログデータの第三の特性は、その取得が測定のための追加的な行為を伴わない点にある。従来の効果検証では、業務時間の実測や担当者へのヒアリング、アンケート調査といった手法が広く用いられてきたが、これらはいずれも観測対象となる職員による記録作業や回答行為を必要とする。その結果、測定そのものが新たな業務負担となり、測定の継続性や現場での受容性を損なう要因となってきた（第 3.1.1 節参照）。

これに対し、デジタル化された行政業務においては、業務の遂行過程における操作や状態変化がシステム上に履歴として自動的に記録されるため、個別の測定行為を実施することなく観測可能なデータを取得することができる。すなわち、業務の実行そのものが、計測の前提となるデータ生成過程を内包しており、観測主体の介入を必要としない。

このような測定行為を伴わないデータ取得は、従来の効果検証手法に内在していた観測に伴う負担の問題を構造的に軽減し得るだけでなく、職員の行動変容を引き起こす観測バイアスの影響を緩和する可能性を有している。実測やアンケート等においては、測定対象者が観測されていることを認識することで、業務遂行の方法や速度に変化が生じ得るが、ログデータは日常の業務運用の過程において副次的に生成されるため、このような影響を回避し得る。

さらに、測定行為を伴わない観測は、評価の対象となる個別業務に対する直接的な介入を必要としないため、施策の対象となった業務以外の工程や部門に生じる影響についても把握可能とする。たとえば、非コア業務に対する自動化施策の導入が、後続の審査業務や意思決定工程といったコア業務の処理時間や滞留に間接的な影響を及ぼす場合であっても、ログに基づく観測はこれらの変化をプロセス横断的に捉えることを可能とする。したがって、ログを用いたデジタル測定は、単一業務内の改善効果にとどまらず、行政サービス提供プロセス全体に対する波及効果を把握するための土台となり得る。このようなプロセス横断的な観測可能性は、施策評価の対象を個別業務の効率化から行政活動全体の価値創出へと拡張するうえで重要な意義を有する。

3.2.5 デジタル測定が内在するリスク

ログデータの活用は、行政活動を低負荷かつ継続的に観測可能とする一方で、プライバシー保護や組織運営の観点から新たな制度的・倫理的リスクを伴う可能性がある。

第一に、業務の遂行過程が履歴として可視化されることにより、個々の職員の行動や活動量が間接的に把握され得る状況が生まれる。このようなデータの利用が、業務改善やサービス品質向上といった本来の目的から逸脱し、個人評価や監督目的へと転用される場合には、職員の自律的な業務遂行を阻害する「監視型管理」への懸念が生じ得る。

第二に、測定の容易化は、測定そのものの目的化を招くリスクを孕む。すなわち、本来は業務改善の手段として導入されたはずの効果検証が、新たな管理指標の達成を目的とした運用へと転化し、行政現場における「測りすぎ」を助長する可能性がある。このような状況は、測定結果の解釈が形式的な数値評価に偏重することを通じて、行政サービスの実質的価値の把握をむしろ困難にする要因となり得る。

第三に、来庁履歴や移動履歴等のログは、個人の行動パターンや勤務実態に関する情報を内包し得るため、適切な匿名化や目的限定の措置を伴わない場合には、プライバシー侵害のリスクが生じる。特に、長期的な継続観測の実施にあたっては、データの取得範囲や利用目的を明確化し、業務改善以外への転用を制度的に制約する必要がある。

したがって、ログを活用したデジタル測定の導入にあたっては、単に計測精度の向上を志向するのではなく、「監視にならない測り方」を前提とした運用原則の設計が不可欠である。本研究では、このような観点を踏まえたうえで、分析対象となるログの範囲および利用目的を第4章において明示的に限定する。

3.3 ログによって可能になる効果検証の拡張

ログデータの活用は、単一業務における処理時間や作業量の変化にとどまらず、業務間の相互依存関係を通じて間接的に生じる影響の把握を可能とする点でも重要な意味を持つ。行政業務、特に内部業務（G2E）は、複数部署にまたがる工程連鎖の中で遂行されることが多く、ある業務プロセスにおける改善が、他業務の処理負荷や滞留時間に影響を及ぼすことがある。

従来の効果検証では、施策の対象となった業務プロセス自体の改善効果に着目する傾向が強く、こうした部門間の波及効果を定量的に把握することは困難であった。その結果、非コア業務に対する自動化施策の導入などにより後続の審査業務や意思決定工程に間接的な影響が生じた場合であっても、評価の対象外として扱われることが多かった。

しかし、ログデータに基づくプロセス横断的な観測が可能となることで、業務Aの改善が業務Bの処理時間や作業量、さらには差戻し回数や再提出率といった処理品質に与えた影響を時系列的に追跡することが可能となる。このような間接効果の可視化は、施策の評価単位を対象業務に限定せず、行政サービス提供プロセス全体へと拡張するものである。

特に、非コア業務の効率化が後続のコア業務に時間的余裕をもたらし、審査判断の一貫

性向上や手戻りの減少といった処理品質の改善につながる場合、それは単一業務内での時間削減以上に行政価値の創出として評価され得る。ログデータを用いた観測は、このような効率性と品質の双方にまたがる波及効果を把握することを可能とし、施策評価の対象を個別業務の改善から行政活動全体の価値創出へと拡張する土台となり得る。

3.3.1 前後、自治体間の反実仮想分析

効果検証においては、施策の導入によって観測された変化が、当該施策そのものによる影響であったのか、それとも外的要因による変動であったのかを識別することが重要な論点となる。このような識別は、しばしば「介入が行われなかった場合に観測されたであろう結果（反実仮想）」との比較によって行われる（第1章参照）。

しかし従来の行政業務においては、導入前の基準値が取得されていない、あるいは年度単位などの粗い集計データしか存在しないことが多く、施策導入の効果を前後比較によって検証すること自体が困難であった（第3.1節参照）。また、業務量の増減や制度改正、需要の季節変動といった外的要因も同時に変化するため、施策導入前後の単純な差分から因果的な効果を抽出することは容易ではない。

これに対し、ログデータに基づく業務観測が可能となった場合、施策導入前後における業務処理の履歴を案件単位で継続的に取得することができるため、短期的な断面比較ではなく、時系列的な変化を追跡することが可能となる。また、行政業務の多くは法令根拠に基づく手続であり、その基本的な処理フローは自治体間で一定の制度・業務構造の共通性を有していることから、類似業務を対象とした自治体間比較を設計することも可能となる。

このような導入前後の継続的な観測および自治体間比較の組み合わせは、施策が導入された群と導入されていない群との変化を比較する準実験的な分析設計を可能とし、行政DXによる介入効果をより厳密に識別するための反実仮想分析に向けた前提条件を整えることを可能とする。したがって、ログデータの活用は、従来困難であった行政施策の因果的解釈に必要な観測条件の整備に対して、新たな分析可能性を提供する土台となり得る。

3.3.2 工程分析・ボトルネック可視化

従来は、紙媒体を土台とした業務運用のもとでは、こうした工程ごとの処理状況を定量的に把握することが困難であり、業務全体のリードタイムがどの工程において発生しているのかを特定することは容易ではなかった。その結果、業務改善の対象は経験的な判断に基づき選定されることが多く、必ずしも処理遅延の主因に対する介入が行われているとは限らなかった。

これに対し、ログデータに基づくプロセス横断的な観測が可能となることで、案件単位の処理履歴をもとに各工程における開始時刻、終了時刻、待機時間などを復元することができる。これにより、業務プロセス内における滞留時間の分布や工程間の引継ぎ頻度を把握し、サービス提供全体の所要時間に対して影響の大きい工程、すなわちボトルネックとなっている処理段階を定量的に特定することが可能となる。

また、こうした工程分析は、単一業務の効率化にとどまらず、前段階の処理遅延が後続工程の処理負荷に与える影響を評価することにも資する。たとえば、申請内容の確認工程における滞留が審査工程の開始時刻を遅延させている場合、審査工程そのものの効率性を改善するよりも、前段階における入力支援や書類確認プロセスの改善が全体の処理時間短縮に寄与する可能性がある。

このように、ログデータを用いた工程分析は、行政業務の実行過程をプロセス単位で可視化することで、改善対象となる工程の優先順位付けを可能とし、業務改善施策の効果的な設計に資する情報を提供するものである。したがって、ボトルネックの所在をデータに基づいて特定することは、行政サービス提供プロセス全体の効率性向上および品質改善を図るうえで重要な意義を有する。

3.4 デジタル測定を活用した評価の実例

ログデータを活用したデジタル測定の有効性は、理論的な可能性にとどまらず、すでに一部の自治体や行政機関において実務的に導入が進みつつある。特に、業務システム上に記録された処理履歴を用いて業務プロセスを可視化するプロセスマイニングの手法は、業務改善の効果検証において活用され始めている。

従来の業務改善は、担当者へのヒアリングや業務フロー図に基づく分析など、主として定性的な情報に依拠して実施されることが多かった。しかし、デジタル化された業務環境においては、案件単位の処理履歴がログとして記録されることで、実際にどの工程で滞留が生じているのか、あるいはどの部署間の引継ぎに時間を要しているのかといった業務プロセス上のボトルネックを、定量的に把握することが可能となる。

たとえば、複数の自治体においては、内部事務処理や住民対応業務の履歴データを用いて、申請受付から決裁までの業務フローを再構築し、工程別の所要時間や待機時間の分布を分析する取組が報告されている。これにより、従来は経験的に推定されていた業務上の負担箇所を実測データに基づいて特定し、処理順序の見直しや業務分担の再設計につなげることが可能となった。

また、こうしたログベースの分析は、改善施策の導入前後における業務プロセスの変化を継続的に追跡することを可能とするため、単なる一時的な業務負担の軽減にとどまらず、改善施策が庁内の業務連携に与えた影響を長期的に評価する手段としても機能し得る。すなわち、デジタル測定は、業務の個別改善に対する効果検証を超えて、行政サービ

ス提供プロセス全体の運用状況を可視化するための継続的な評価に活用する枠組みとして位置づけられる。

以上のように、ログデータを活用したデジタル測定は、行政業務の工程単位での実行状況を客観的かつ継続的に観測可能とすることで、従来は困難であった業務プロセスの改善効果の把握や、組織内外における業務連携の変化の評価に資する情報を提供するものである。

3.4.1 群馬県前橋市・高崎市・伊勢崎市

前橋市・高崎市・伊勢崎市では、住民基本台帳業務において、転入・転出処理や証明書発行、マイナンバーカード関連業務などが職員の手作業を前提に遂行されており、業務量の増加や属人化、改善余地の把握が困難であることが課題となっていた。特に、業務実態の把握はマニュアルやヒアリングに依存しており、実際の処理手順や例外対応、処理時間のばらつきが十分に可視化されていない状況にあった。

この課題に対し、住民基本台帳システムのアクセスログを活用したプロセスマイニングに着手し、業務の実際の流れを客観的に把握する取組を実施した。アクセスログから画面遷移や操作履歴を抽出し、標準的な処理パターンを「トランザクション（1 手続き単位）」として定義することで、転入 1 件あたりの処理時間や工程別の所要時間を定量的に把握できる分析基盤を構築した。

分析にあたっては、サンプリングではなく全件ログを対象とし、市別・庁舎別・職員別の処理時間や業務フローの違いを可視化した。その結果、同一業務であっても手順や処理順の違いによって生産性や修正頻度に差があること、慣例的に実施されていた一部作業が法令上は不要であることなどが明らかとなった。これにより、単なる自動化候補の抽出にとどまらず、業務そのものの廃止・簡素化や標準化の判断が可能となった。

また、アクセスログ分析で得られた定量的データをもとに、RPA 適用が有効な定型・反復業務を選定し、業務プロセス改善と自動化を段階的に検討した。効果の評価は、「何件処理したか」ではなく「1 手続きあたりの処理時間」を主な指標とし、業務改善効果を客観的に測定できる枠組みを採用した。

この取組により、アクセスログを監査用途にとどまらず、業務改善・RPA 検討・人材育成・ガバナンス強化に活用する可能性が示され、複数自治体で横展開可能な業務プロセス改善モデルとして整理された。

3.4.2 カナダ政府

カナダでは、プロセスマイニングを行政業務の可視化・改善に活用するための実践的な枠組みとして、PM4Canada Process Mining Framework が整備されている（PM4Canada,

n.d.)。同フレームワークは、単にログデータを分析する手順を示すにとどまらず、分析対象とするプロセスの選定 (Process Qualification)、データの整備 (Data Preparation)、分析の実施 (Analysis) という三段階の構造を採り、とりわけ分析に着手する前段階（すなわち「何を測る価値があるか」「必要なデータにアクセスできるか」を組織的に判断する工程）に重心を置いている点に特徴がある。

第一段階の Process Qualification では、対象とする業務プロセスの選定、研究課題の定義、およびステークホルダーとの合意形成を行う。プロセスマイニングの失敗原因の多くがこの段階の不備に起因するとされ、分析の投資対効果 (ROI) が見込める案件のみを選別する仕組みが設計されている。第二段階の Data Preparation では、対象業務に関するデータの特定、アクセス権の交渉、データクレンジング、およびイベントログへの変換を行う。フレームワークが強調するのは、プロジェクト全体の工数の大半がこの段階に費やされるという実態であり、イベントログの最低要件として、案件 ID (case_id)、イベント種別 (event)、タイムスタンプ (timestamp) の三項目が定義されている。第三段階の Analysis では、構築されたイベントログに基づき、プロセスモデルの生成、ボトルネックの特定、処理時間の分布分析、SLA との乖離の把握等を行う。

このフレームワークの意義は、技術的な分析手法の提供にとどまらず、公共セクター特有の障壁であるデータ共有への組織的抵抗、業務可視化に対する心理的懸念、調達や権限管理の制約を前提として設計されている点にある。たとえば、組織内での採用を促進するためのブリーフィングノートの雛形が用意されており、合意形成プロセスが制度化されている。これは、本研究が第 2 章および第 3.1 節で整理した「技術的にはログが存在しても、組織・運用の条件が整わなければ測定が実装されない」という構造的課題に対する、実務的な応答として位置づけることができる。

同フレームワークの実証的適用として、オタワ大学の研究グループは、カナダ連邦政府における職員セキュリティ審査 (Personnel Security Screening) プロセスを対象に、プロセスマイニングによる業務分析と改善効果の検証を実施した (Latheef, 2024)。同研究は、改善介入の前後において同一プロセスのイベントログを分析し、業務プロセスの構造的変化と処理効率の改善を定量的に把握した点で、行政業務におけるプロセスマイニングの有効性を示す事例として注目される。

分析の結果、改善前のプロセスには、例外処理（修正案件）の頻発、手戻り（ループ）の多発、本来関与すべきでない担当者の介入といった構造的な非効率が検出された。これらはいずれも、ヒアリングや業務フロー図だけでは定量的に把握しにくい問題であり、イベントログに基づく可視化によってはじめて明確化されたものである。

改善介入として、低コードのデジタルプラットフォームを活用したプロセスの再設計およびリソース配置の見直しが実施された。介入後に再度イベントログを分析した結果、セキュリティブリーフィングに要する時間が約 7 日から約 46 時間へ短縮され、プロセス全体の処理期間も約 31 日から約 26 日へ改善された。また、一案件あたりの関与人数が削減

され、人的コストの低減も確認された。

本事例の本質的な意義は、数値の改善幅そのものよりも、「可視化 → 改善 → 再測定」という循環が、同一のイベントログを基盤として一貫して実現されている点にある。すなわち、改善前の分析（pre）でボトルネックを特定し、改善介入を設計・実施したうえで、改善後の分析（post）で効果を検証するという、ログを起点とした改善ループが成立している。これは、本研究が第 3.2 節で論じた「ログの自動性・全件性・連続性」が、単なる現状把握の道具にとどまらず、改善施策の因果的効果を事前・事後で検証する評価に活用する枠組みとして機能し得ることを実証的に示すものである。

また、効果検証に用いられた指標が処理時間と関与人数（人件費の代理指標）というシンプルな構成であった点も示唆的である。これは、イベントログから直接算出可能な指標のみで、実務上十分な改善根拠を提供できることを意味しており、本研究が提案する「ログで測れるものから固める」というモデル設計の基本方針（第 4.1 節）と整合する。

一方で、同フレームワークおよび事例研究には、KPI 設計（何を測るべきか）の体系的な方法論、政策効果との接続、定性的評価との統合といった論点が明示的には含まれていない。したがって、プロセスマイニングは業務プロセスの可視化と改善効果の検証において強力な手法であるが、行政 DX の成果を多面的に把握するためには、本研究が第 1 章で整理した指標体系や評価フレームワークとの接続が別途必要となる。

3.4.3 静岡県裾野市

裾野市では、窓口業務の実態把握が長年の課題であった。どの手続にどれくらいの時間を要しているか、どの時間帯にどの程度の来庁があるかといった基本的な業務構造が、現場の体感や経験則に依存しており、客観的な根拠に基づく改善判断が困難な状況にあった。個別のタイムスタディ⁷⁷による計測は職員への負担が大きく継続が難しいうえ、手入力による記録は実務との乖離が生じやすく精度にも課題があった。

この状況を打開する契機となったのが、「自然に取得されるデータ」への着目である。裾野市が導入した窓口予約・発券システム（フロントデスク）は、受付・呼出・処理完了の各時点をログとして自動的に記録する。日常業務の遂行と同時にデータが蓄積される設計であり、職員に追加的な入力作業をほぼ求めない。裾野市は、この特性に着目し、「業務を止めずに取得できるログ」を基盤として窓口業務を把握する方針へ転換した。

具体的には、市民課窓口に限定して導入されていた同システムを、庁舎 1 階フロアにおける窓口業務全体へ拡張した。この拡張により、窓口対応、電話対応、バックオフィス業務といった業務区分を横断して、手続処理データを一体的に取得・集計することが可能と

⁷⁷ タイムスタディ（Time Study）：ストップウォッチ等で作業工程を直接計測し、標準所要時間やボトルネックを把握する作業測定法。

なった。なお、この拡張に先立ち、裾野市では業務プロセスそのものの見直し（BPR）が徹底的に実施されており、処理時間は極限まで短縮された状態にある。その結果、フロントデスクのログで把握できる窓口対応時間が、内部処理時間をも概ね反映する構造となっている点は、ログの測定射程を考えるうえで重要である。

蓄積されたログは、日々確認可能なダッシュボードとして整備されている。職員は自らの業務の状態を迅速に把握できる環境のもと、業務の状態を「体温計のように」継続的に確認しながら、日常業務の中で小さな改善を積み重ねている。ダッシュボードの内容自体も固定的ではなく、業務の実態に即して見直されながら運用されており、日次レベルでの観測と改善が接続された継続的な PDCA サイクルが形成されている。

データ活用の具体例として、時間帯ごとの来庁ピークの把握による人員配置の見直し、手続ごとの処理時間のばらつき分析に基づく手続の統合・分割の検討、特定時間帯や曜日における待ち時間の異常値検知と原因分析が挙げられる。こうしたデータに基づく意思決定により、従来は体感に依存していた改善判断が客観化され、改善の優先順位が明確になるとともに、対内部・対外の説明責任にも活用可能な根拠が整備されている。

一方で、導入当初は、処理時間の可視化に対して「個人の能力が評価されるのではないか」という心理的懸念が職員の間に存在した。これに対し、データの目的が個人評価ではなく業務改善にあることを組織として明確化し、現場全体の改善に用いるという方針を共有することで、徐々に受容が進んだ。その結果、データを見る文化が現場に浸透し、職員側からも改善提案が出るようになっていく。この過程は、ログ活用においては技術的な仕組みの整備だけでなく、データをどう使うか、現場が納得して使えるかという組織面の設計が定着の鍵を握ることを示している。

裾野市は、現在の窓口業務にとどまらず、他業務への横展開や、蓄積されたデータを用いた因果分析・効果検証への発展も視野に入れている。同システムは座間市や酒田市等にも導入されており、同一の窓口予約・発券システムを用いる自治体間で共通のログ構造に基づく比較分析が可能になる点は、自治体横断のベンチマーキングの観点からも注目される。

このように、裾野市の事例は、「測る仕組み」を新たに構築するのではなく、「業務の過程で自然に貯まる仕組み」を土台として業務改善と効果検証を接続するモデルとして位置づけることができる。負担を増やす計測は定着しないという実務上の制約を前提に、既存のシステムログを起点として、可視化・改善・受容・横展開を段階的に進めてきた本事例は、本研究が提案する「追加負荷ゼロないし極小で測れる範囲はどこまでか」という測定設計の基本方針と、その実装可能性を具体的に示すものである。

3.4.4 東京都練馬区

練馬区では、区民部収納課における未納対応業務において、財産調査先の選定や財産

の特定が担当職員の経験に依存しやすく、ベテラン職員と新人職員の間で調査効率に差が生じていた。この課題に対し、ベテラン職員の判断ロジックを AI に学習させることで、財産調査先の推定を自動化・高度化する取組を実施した。効果検証では、処理時間をシステム操作履歴（ログ）から取得するとともに、差押え件数・徴収額は毎年度の業務データを用いて導入前後を比較する方式を採用した。その結果、財産調査業務において 1 件あたり約 30 分を要していた作業が約 3 分に短縮されたほか、財産特定の成功率が向上し、差押え件数および徴収額の増加が確認された。加えて、AI がベテランのノウハウを再現することで新人職員でも同等の成果を出せる状態が実現され、属人性⁸の解消にも寄与している。さらに、業務効率化により創出された余剰時間は、納付困難な住民への福祉的支援（困窮者対応・納付相談等）に充てられるようになり、単なる時間短縮にとどまらず、対応の質的転換にもつながっている。

3.4.5 民間事例

民間のプロセスマイニング事例として、日立システムズは、業務システムのイベントログ等に基づき実際の業務プロセスを可視化・分析することで、RPA 等の部分最適では見えにくいボトルネックや課題をファクトデータとして把握し、改善につなげる取組を紹介している。経験則で「課題」と見なしていた点が、データ上は必ずしも主要因ではないと判明する場合もあるため、分析を反復しながら課題設定を更新する姿勢が重視される。また、継続運用の前提として「測れるデータ構造（ログ設計・タイムスタンプ等）」の整備と、全体最適の視点での継続的な分析・改善サイクルの確立が要諦とされる。

3.5 小括

第一に、3.1 節では、従来の効果検証が困難であった理由を、計測そのものが現場に追加的負担を生じさせること、効果検証に耐えうる観測時点および観測粒度が不足していること、さらに長期的・継続的な測定の設計と運用が困難であること、の三点から整理した。実測、ヒアリング、アンケート等の手法は、業務実態を把握するうえで一定の有効性を有する一方、測定のために現場へ新たな観察・記録作業を課すため、負担と継続性との間に構造的なトレードオフを抱えていた。

第二に、3.2 節では、こうした困難に対し、業務遂行の過程で自動的に生成されるログが、有力な測定の仕組みとなりうることを示した。ログは、測定行為を伴うことなく取得できるため、現場に追加的な負担を生じさせにくい。また、特定時点のサンプルではな

⁸ ここでいう属人性とは、業務の遂行や判断が、標準化された手順や共有されたデータではなく、特定職員の経験・ノウハウに依存している状態を指す。

く、案件単位の全件データを連続的に把握しうることから、処理時間、待機時間、差戻し回数、機能利用状況といった比較的客観性の高い対象を継続的に観測することが可能となる。他方で、端末操作履歴等の活用には、監視として受け取られるリスクが内在しているため、取得目的の限定、匿名化・仮名化、アクセス統制等の運用上の規律を前提とする必要があることも確認した。

第三に、3.3 節では、ログを用いることで、単なる現状把握にとどまらず、前後比較や自治体間比較といった反実仮想的な分析、および工程分析やボトルネックの可視化が可能となることを示した。さらに 3.4 節では、高崎市・前橋市・伊勢崎市、カナダ政府、裾野市、練馬区、民間における活用例を通じて、ログを土台とする測定が、実務において成果の可視化と改善の循環に貢献しうることを確認した。これらの事例は、ログが単なる記録ではなく、改善前後を比較し、業務プロセスの変化を継続的に捉える評価に活用する枠組みとなりうることを示している。

以上の整理から、本章で明らかになったのは、ログを中心とするデジタル測定が、従来の効果検証に内在していた負担、観測時点、継続性の制約を緩和し、低負荷で継続可能な評価に活用する枠組みを構築する可能性を有するという点である。ただし、ログは万能ではなく、ログによって安定的に取得できる範囲と、取得が困難な範囲とを区別したうえで、観測困難な領域については実測、ヒアリング、アンケート等によって補完する設計が必要である。

したがって、次章で求められるのは、ログ活用の可能性を理念として示すことではなく、観測可能なデータと非観測データを峻別し、業務ログに基づく処理時間・滞留状況の把握、導入ステップ、活用ケースを含むかたちで、実務上導入可能な「ログを用いた低負荷・継続的なデジタル測定による効果検証モデル」として具体化することである。第 4 章では、本章で整理した知見を踏まえ、低負荷かつ継続運用可能なモデルの構造と運用条件を提示する。

4. ログを活用し、業務と測定を一体化する行政 DX 効果検証モデル

の検討

本章では、前章までの論点を踏まえ、ログデータを中心とした「ログを用いた低負荷・継続的なデジタル測定による効果検証モデル」を提案する。モデルは、(1) 観測可能データと非観測データの整理、(2) 業務ログを用いた処理時間・滞留状況の把握、(3) 工程別のボトルネック特定、(4) 継続運用可能な年次サイクルの構築、という 4 つの要素で構成する。本章の目的は、行政 DX の成果を追加負担なく継続的に可視化できる仕組みを設計し、実務で導入可能な評価体系として提示することである。また、小さく始めて徐々に改善するための導入ステップや運用ガイドラインも整理し、自治体が現実的に取り組める枠組みを示す。

4.1 モデル設計の基本方針

本章で提案する行政 DX 効果検証モデルは、前章までで整理してきた課題認識を踏まえ、行政現場に追加的な測定負担を生じさせることなく、継続的かつ比較可能な形で成果を把握することを目的とする。従来の効果検証は、実測、ヒアリング、アンケート等に依存することで、測定そのものが新たな業務負担となりやすく、また観測時点や観測粒度の不足、年度をまたぐ継続運用の難しさといった制約を抱えていた。これに対し、行政 DX の進展に伴って蓄積される業務ログは、業務遂行の過程で自動的に生成され、案件単位で全件的・連続的に取得できる可能性を有している。したがって、今後の効果検証においては、ログを中心に据えた観測設計を行うことで、従来手法の限界を補いながら、低負荷で持続的な評価に活用する枠組みを構築することが重要となる。

ここで、本モデルの射程を明確にしておく。本モデルが直接の観測対象とするのは、業務ログから安定的に取得可能なプロセス指標（処理時間、待機時間、差戻し回数、操作回数、引継ぎ頻度等）およびアウトプット指標（処理件数等）である。すなわち、本モデルは「業務がどのように遂行されたか」を継続観測する層に特化しており、「その結果として住民・職員・組織にどのような価値変化が生じたか」というアウトカム層を直接測定するものではない。来庁回数の削減や手続完了までの所要日数の短縮は、住民利便性の向上を示唆する代理指標として位置づけ得るが、住民の満足度・安心感、職員の心理的負担の軽減といったアウトカムそのものとは区別される必要がある。本モデルは、プロセス・アウトプット層の観測を安定化させることで、アウトカム評価の前提となるファクトベースを整備するものであり、アウトカム評価そのものは補完的手法との組み合わせによって実現される（4.5 節参照）。

もっとも、行政業務のすべてがログのみで完全に観測できるわけではない。実際には、

机上審査、電話確認、窓口での口頭対応、個別事情への判断といったシステム外の工程が一定程度残存しており、また記録粒度やタイムスタンプ定義の不統一、連結キーの欠落、ログ欠落や例外処理の存在も、観測の精度と比較可能性を損なう要因となる。そのため、本モデルでは、ログを万能なデータ源として扱うのではなく、まず観測可能なデータと非観測データを峻別したうえで、ログによって安定的に把握できる部分を評価の基盤とし、必要に応じて補正や補完の仕組みを組み合わせるという設計思想を採用する。

さらに、本モデルは単に処理時間や件数の変化を機械的に集計することを目的としない。行政 DX の成果は、住民・事業者の利便性、職員負担の軽減、組織運営の効率化、自治体間連携の円滑化といった多面的な価値として現れるため、時間・費用・利便性・職員負担・組織運営といった複数の指標体系との接続を意識した設計が必要である。したがって、本章で提案するモデルは、ログを用いて観測可能な業務実行データを基礎に置きつつ、それを行政価値の評価へと接続する「ログを用いた測定体系」として構成する。

この基本方針のもと、本節では第一に、観測対象となるデータを「観測データ」と「非観測データ」に区分し、評価可能範囲を明確化する。第二に、業務ログを中心とした「ログを用いた測定体系」の考え方を示し、追加的な調査負担を最小化しつつ継続観測を実現する方法を整理する。第三に、こうした測定を単発の実証に終わらせず、年度をまたいで安定的に運用可能とするための制度・運用上の要件を提示する。

4.1.1 観測データと非観測データの整理

行政 DX の効果検証モデルを設計するうえでは、まず、何がデータとして観測可能であり、何が観測困難であるのかを整理する必要がある。行政業務のデジタル化が進むことで、申請受付、審査開始、差戻し、承認、完了といった処理工程は、電子申請システムやワークフロー等を通じて履歴として記録されるようになっている。本研究では、このように業務の遂行と同時に自動的に生成される履歴情報を「観測データ」と位置づける。

一方で、行政業務には、システム上に十分に記録されない活動や、ログのみからは把握しにくい価値も存在する。たとえば、窓口での説明、電話やメールによる照会対応、個別事情を踏まえた判断などは、業務全体に影響を与えるが、ログとして直接取得できない場合が多い。また、住民や事業者の満足度、職員の心理的負担、組織内の納得感といった価値も、ログのみから把握することは難しい。本研究では、このような情報を「非観測データ」と整理する。

非観測データの代表例としては、窓口対応や電話・メールでの説明、個別判断を要する例外処理などに加え、職員が業務遂行において負う心理的負担やストレスのように、ログには痕跡として残らないが業務の質や持続可能性に影響する要素が挙げられる。

重要なのは、観測可能なデータのみが重要であるということではない。行政 DX の成果は本来、多面的な価値として現れるため、ログで把握できる部分と、ログでは把握しにく

い部分の双方を視野に入れる必要がある。そのうえで本モデルでは、継続的かつ低負荷に取得できるログを評価の土台としつつ、ログだけでは捉えきれない価値については別途補完の必要があるという整理を採用する。

4.1.2 業務ログを中心とした「ログを用いた測定体系」

行政 DX の効果検証を実務的に成立させるためには、評価指標や分析手法の設計以前に、「どのようなデータが、どの程度の負担で、継続的に取得可能か」をあらかじめ設計することが不可欠である。本研究では、この測定設計を二つのレイヤーに分けて整理する。

第一に、業務遂行の副産物として自動的に生成されるログやメタデータを活用した、受動的かつ全件的・継続的に把握する仕組みの構築である。申請システム等の明示的な業務システムに加え、コミュニケーション基盤、端末利用、ネットワーク、物理的入退室といった周辺インフラを含めた観測設計を行うことで、現場に測定負担を課すことなく業務実態を把握することが可能となる。

第二に、これらの観測データから、直接取得できない業務時間や調整負荷といった非観測量を、論理的・統計的に推定するアプローチである。特に、DX 導入以前の業務実態については過去データが存在しない場合も多いが、導入後のログとマクロな勤怠情報を組み合わせることで、一定の仮定の下で事前状態を推定することが可能となる。

このような測定体系は、単発の事後評価にとどまらず、低負荷で継続的な観測を可能とする点に特徴があり、結果として事後評価で得られた知見を次の施策設計（事前評価）に還元するサイクルの土台となる。

4.1.3 測定負担を最小化し、継続運用可能な仕組み

業務ログを活用した測定体系の最大の利点は、効果検証そのものが現場の新たな負担とならない点にある。従来の効果検証では、実測、ヒアリング、アンケート等の手法が多用され、測定行為そのものが職員の業務時間を消費し、結果として測定の継続性を損なってきた。本研究が提案するモデルでは、測定の主軸を業務ログに置くことで、こうした負担構造を根本から転換する。

また、一次情報の入手に係る工数を削減できることは、単に調査負担を軽減するという消極的な利点にとどまらない。原課担当者が情報提供や説明対応に過度な時間を割かずに済む分、解析結果の解釈や改善策に関するディスカッションに、より多く参画できるようになる。すなわち、測定負担の削減は、データ収集の効率化にとどまらず、改善活動そのものの質と参加度を高める条件として位置づけられる。

4.2 業務ログに基づく処理時間・滞留状況の把握モデル

本節では、業務システム等に記録されたログデータを用いて、案件単位・工程単位の所要時間を推定するモデルを整理する。あわせて、ログでは直接観測できない工程については、短期実測や補完的情報を用いた補正の考え方を示し、実務上利用可能な推定枠組みを提示する。

4.2.1 業務単位の処理時間・滞留状況の把握

業務単位の処理時間・滞留状況の把握では、受付、審査開始、差戻し、決裁、完了といったイベント時刻をイベント ID で連結し、業務全体のリードタイムおよび工程別の処理時間を算出する。これにより、従来は把握が難しかった案件単位の処理実態を、追加的な測定負担なく継続的に観測することが可能となる。具体的には、第一に、受付から完了までの総所要時間を把握し、業務全体のリードタイムを確認する。第二に、受付、審査、確認、決裁、通知等の工程別所要時間を算出し、遅延がどの工程に生じているのかを把握する。第三に、工程間の待機時間を把握し、処理時間と滞留時間を区別する。これにより、単に業務全体が長いかわかりだけでなく、どの工程で処理が行われ、どの工程間で待ちが発生しているのかを把握できる。

ただし、窓口での口頭説明、電話確認、紙資料の確認、個別判断、現地確認など、システム外で行われる工程はログに直接残らない場合がある。そのため、ログから算出される所要時間は観測可能な工程に基づく値として扱い、必要に応じて短期実測、ヒアリング、業務日誌等で補完する。補完値はログから直接得られる観測値とは区別し、推定条件や対象範囲を明記する必要がある。

4.2.2 ボトルネック特定手法（待ち時間／操作回数）

ボトルネックの特定にあたっては、工程間の待機時間、差戻し回数、再処理頻度、操作回数等を主要な観測指標として用いる。これにより、処理遅延の主因が個別工程の処理能力にあるのか、前段階の確認不足や引継ぎの滞留にあるのかを、プロセス横断的に把握することができる。

4.3 導入ステップ（運用モデル）

ログを活用した効果検証モデルを実務に導入するためには、分析手法の高度化に先立

ち、継続的に観測できる最小限のデータ構造と、それを運用するための体制を整える必要がある。特に、行政現場では、対象業務の範囲、システムごとの記録粒度、データへのアクセス権限、人事異動や年度予算の制約が複雑に関係するため、最初から全庁的・網羅的な測定 of の仕組みを構築することは現実的ではない。むしろ、業務量が多く、状態変化が明確で、既存システム上に一定の履歴が残っている業務を対象として、小さく観測を始め、分析と改善の循環が成立することを確認しながら段階的に拡張していくことが重要である。

本節では、こうした考え方にに基づき、導入ステップを三つの観点から整理する。第一に、案件単位・工程単位の処理実態を復元するために必要となるログ収集の最低要件を示す。第二に、取得したログを一度きりの分析に終わらず、観測、分析、改善、再観測を繰り返す年次サイクルとして制度化する方法を整理する。第三に、PoC から開始し、対象工程や対象部門を段階的に拡張しながら、監視や個人評価への転用を避けるためのガードレール⁹を組み込む運用上の留意点を示す。

これにより、本モデルは単なる分析手法ではなく、自治体の実務の中で継続的に運用し得る効果検証の仕組みとして位置づけられる。

4.3.1 ログ収集の最低要件

行政 DX の効果検証を実務において成立させるためには、まず、分析の高度さ以前に、継続的な観測を可能にする最低限のログ要件を整備する必要がある。とりわけ重要なのは、案件を一意に識別する case、処理や状態変化を表す activity、当該処理がいつ発生したかを示す timestamp、および担当組織または処理主体を示す resource の四要素である。これらは、案件単位・工程単位での処理実態を復元し、所要時間、待機時間、差戻し、再処理といった指標を安定的に算出するための最小スキーマとして位置づけられる。

ただし、必要なのは単に項目を記録することではない。継続観測と比較可能性を担保するためには、case の単位を何に置くのか、activity をどの粒度で定義するのか、timestamp をどの状態変化に付与するのか、といった定義の最低標準化が不可欠である。同じ「受付」や「完了」という名称であっても、部局や年度によって意味内容が揺らげば、前後比較や横断比較の基盤は失われる。もっとも、法令根拠が共通する業務であっても、実際には自治体や部局ごとに業務フローや運用に差異があるため、厳密な意味での完全統一は現実的でない。ここで求められるのは、絶対的に同一な測定ルールではなく、観測結果を相対比較可能な水準へと引き上げるための最小限の定義・記録仕様である。このため、ログ項目そのものに加え、語彙、記録粒度、連結キー、タイムスタンプ定義をあらかじめ揃え

⁹ ガードレールとは、測定やデータ活用が過度な負担や不適切な利用に至らないようにするための制約や運用上のルールを指す。

ておくことが、実務上の最低条件となる。

さらに、最低要件とは「すべてを最初から取得すること」を意味しない。むしろ重要なのは、後続の分析と改善に必要な最小限の構造を確保し、運用可能な範囲から着手することである。最初の段階では、全業務を対象とするのではなく、業務量が多く、状態変化が比較的明確で、効果検証の必要性が高い対象に限定して、最小スキーマの整備を優先することが適切である。その際には、必要なログがどの情報システムに存在し、誰がどの権限でアクセスできるのかというデータの所在と権限を併せて確認しておく必要がある。新たな計測の仕組みを構築する前に、既存の業務システムや監査ログ等、すでに蓄積されている記録を起点とすることで、追加的な負担を抑えながら観測を立ち上げることができる。

4.3.2 データ収集 → 分析 → 改善の年次サイクル

効果検証を単発の事後評価に終わらせないためには、ログの収集を一度きりの分析に用いるのではなく、観測 → 分析 → 改善 → 再観測の循環として制度化する必要がある。本研究では、この循環を四半期単位の継続観測を基礎とする年次サイクルとして位置づける。具体的には、繁忙期・閑散期を跨いだデータ収集を行い、一定期間ごとに処理時間、差戻し回数、待機時間、引継ぎ頻度等を分析し、その結果に基づいて改善施策を実施したうえで、再度同一指標により変化を観測する。この反復によって、可視化が改善に接続され、改善が再測定によって検証される運用が成立する。

この年次サイクルを安定的に運用するうえで重要なのは、ログデータそのものだけではなく、それを支える指標の定義書、データ辞書、可視化仕様、メタデータを継続的に更新・管理することである。人事異動やシステム改修があっても同じ意味で指標を読み続けられるよう、定義の版管理と観測手順の文書化を引継ぎ資産として保持する必要がある。これにより、観測の継続性と比較可能性を損なうことなく、年度をまたいだ評価と改善の循環を安定化できる。

また、この年次サイクルは、個々の自治体内に閉じた改善にとどまる必要はない。共通の定義のもとで業務時間、件数、コスト、差戻し等の指標を蓄積できれば、各自治体は自らの観測結果を他団体の水準と相対化し、改善余地や優先順位をより明確に判断できる。さらに、成功事例のみならず、期待どおりの効果が得られなかった事例や運用上の失敗も共有されることで、他の自治体が試行錯誤を重複することを避けられる。こうした横断データ基盤と共通指標の整備は、個別改善の妥当性を高めると同時に、自治体横断の学習の土台としても機能する。

4.3.3 小さく始めて継続するための運用ガイドライン

ログを活用した効果検証は、構想としては全庁横断・自治体横断に拡張し得るが、実務上は、最初から広範な対象に適用するよりも、小さく始めて継続可能性を確認しながら段階

的に拡張することが重要である。導入初期には、まず限定的な対象業務において PoC として着手し、最小スキーマに基づく観測と分析が成立するかを確認する。この初期段階では、必ずしも新規のログ収集基盤を前提とせず、既存システムに蓄積された記録や、改修前の状態を推計で補ったベースラインを用いることで、低コストに観測を立ち上げることも有効である。その後、対象工程の拡張、関連部門への展開、可視化指標の洗練を経て、日常的な業務運用の中に定着させていくことが望ましい。重要なのは、分析の高度化を急ぐことではなく、測定と改善の循環が実務の中で回る単位を見極めながら、無理なく拡張することである。

同時に、運用の初期段階から明確なガードレールを設けることが不可欠である。すなわち、ログ活用の目的を業務改善とサービス品質向上に限定し、個人の能力評価や監督目的に転用しないことを原則として明文化しなければならない。ログは処理工程や機能利用状況といった比較的客観的な観測対象から段階的に活用すべきであり、個々の職員の評価指標として用いられるべきではない。また、最小限収集、匿名化・仮名化、アクセス統制、監査記録といった規律を併せて整備することで、現場に「監視されるための測定」ではなく「改善のための測定」であるという受容性を確保する必要がある。なお、これらのガードレール（取得目的の限定、個人評価への不転用、最小限収集、匿名化・仮名化、アクセス統制等）の詳細は、5.2.2「監視にならない測り方の導入」において改めて整理する。このように、PoC、拡張、定着という段階的な運用モデルと、監視にしない・個人評価に使わないというガードレールをあらかじめ組み合わせて設計することによって、ログを用いた効果検証は、単発の分析プロジェクトではなく、継続的な業務改善の土台として実装可能となる。さらに、こうした取組を一過性で終わらせず継続させるためには、成功事例を組織内外で共有して改善に取り組む現場へのインセンティブを設計することや、観測・改善のサイクルを事業計画・基本計画に紐づけ、制度的・予算的な裏付けを与えることが有効である。

4.4 デジタル測定モデルの活用ケース

本モデルは、住民向け手続、内部業務、自治体間連携といった異なる行政領域に対して、それぞれ異なる観測対象と評価指標を設定することで適用可能である。本節では、G2C、G2E、G2G の三領域を例に、ログデータからどのような成果を把握できるかを具体化する。

4.4.1 G2C：住民手続の負担軽減の可視化

G2C 領域では、申請から完了までの所要時間、差戻し回数、再申請率、来庁を伴う工程の有無などを把握することで、住民の手続負担軽減を代理的に評価することができる。こ

れにより、オンライン化率のような実施状況指標にとどまらず、手続完了までの実質的な利便性向上を継続的に観測することが可能となる。ただし、これらの指標はプロセス改善の度合いを示すものであり、住民が実際に感じる利便性や満足度の変化を直接把握するものではない。アウトカムの検証には、4.5 節で述べる補完手法との接続が必要となる。

4.4.2 G2E：職員業務の工程分析

G2E 領域では、文書管理、決裁、照会回答等の内部プロセスについて、工程別の滞留時間、引継ぎ頻度、差戻しの発生状況を把握することで、業務負荷の偏在や改善余地を可視化できる。特に、個別工程の改善が全体最適につながっているかを確認するうえで、プロセス横断的な観測が重要となる。ただし、滞留時間や引継ぎ頻度の改善は業務プロセスの効率化を示すものであり、職員の心理的負担の軽減や業務の質的改善を直接測定するものではない。こうしたアウトカムの把握には、4.5 節で述べる補完手法の併用が前提となる。

4.4.3 G2G：連携業務の遅延箇所の特定

G2G 領域では、複数機関・複数部門にまたがるデータ連携の履歴を用いて、どの引継ぎ地点で待機や差戻しが発生しているかを把握することが重要となる。これにより、標準化不足や権限管理、データ形式の不整合に起因する遅延箇所を特定し、自治体間連携の改善対象を明確化することができる。ただし、待機や差戻しの削減は連携プロセスの改善を示す指標であり、連携の質的向上や相互運用性に対する組織的な実感を直接把握するものではない。これらのアウトカムについても、4.5 節の補完手法を通じて検証する必要がある。

4.5 ログ観測と補完手法の接続

本モデルが対象とするプロセス・アウトプット層の継続観測は、行政 DX の効果検証における土台としての役割を果たす。しかし、第 1 章で整理したとおり、行政 DX の成果は住民の利便性、職員の負担感、組織運営の質といった多面的なアウトカムとして現れるため、ログのみで効果検証が完結するわけではない。本節では、ログ観測を起点として、補完手法をどのように接続するか設計枠組みを示す。

4.5.1 ログ指標とアウトカムを接続する仮説設計

ログ指標と行政サービスの質を最終的に規定するアウトカムとの間には、構造的な距離がある。処理時間、待機時間、差戻し回数、操作回数といった指標は、業務プロセスの状態を把握するうえで有効であるが、それらの改善が住民の利便性、職員の負担軽減、組織内の連携改善といったアウトカムにそのまま結びつくとは限らない。したがって、ログ指標の改善をもって、直ちに行政 DX の成果が実現したと結論づけることは適切ではない。

そのため、本モデルでは、ログ指標をアウトカムの直接証拠ではなく、アウトカムに接近するための観測可能な手がかりとして位置づける。具体的には、ログで把握できる変化がどのような経路で住民・職員・組織にとっての価値へ接続するのかを、接続仮説として明示する。たとえば、手続完了までの所要日数の短縮は住民の待ち負担軽減に、差戻し回数の減少は手続の分かりやすさに、工程間待機時間の短縮は職員の時間的余裕に接続する可能性がある。

もっとも、これらは仮説であり、ログだけで検証できるものではない。仮説が実際にアウトカムとして成立しているかを確認するためには、アンケート、ヒアリング、業務日誌、短期実測等の補完手法を組み合わせる必要がある。ログ指標と補完手法の結果を突合し、両者が連動していない場合には、プロセス改善がアウトカムに到達していない可能性を示すものとして、次の改善サイクルの論点とする。

4.5.2 補完手法の実施要件

補完手法を「別途実施する」という宣言にとどめず、運用サイクルの中に組み込むためには、最低限以下の三点を設計しておく必要がある。

(1) 何を測るか：ログが到達しない観測対象の特定 対象業務ごとに、ログで把握可能な指標と、ログでは把握できないが評価上不可欠な項目を整理し、後者を補完手法の対象として明示する。典型的には、住民の満足度・納得感、職員の心理的負担、窓口対応における説明の質的側面、創出された時間の用途などが該当する。この整理は、4.1.1 節で示した観測データと非観測データの区分に基づいて行われるべきであり、対象業務の PoC を開始する段階で、ログで測る項目と補完手法で測る項目を併記した観測設計書として文書化しておくことが望ましい。

(2) いつ測るか：年次サイクルへの組み込み 4.3.2 節で整理した年次サイクル（観測 → 分析 → 改善 → 再観測）の中に、補完手法の実施タイミングをあらかじめ位置づける。ログによる四半期ごとの定量観測に対して、年 1 回から 2 回程度のアンケートまたはヒアリングを組み合わせることで、プロセス指標の改善がアウトカムの変化として現れているかを確認する。補完手法は大規模である必要はなく、対象業務の関係者（住民・職員）に対する少数の設問による定点観測でも、ログとの突合において十分な意味を持ちうる。重要なのは、補完手法が年次サイクルの中に制度的に組み込まれ、毎年度の実施が前提化されていることである。

(3) どう接続するか：ログ指標との突合 補完手法で得られた結果を、同時期のログ指標と突合することが、接続設計の核心である。たとえば、「処理時間は短縮したが住民の満足度は変わらない」「差戻しは減ったが手続の分かりやすさの実感は伴っていない」「操作回数は減ったが職員の負担感は軽減していない」といった乖離が検出される場合、それはプロセス改善がアウトカムに到達していないことを意味する。こうした乖離こそが、次の改善サイクルにおいて優先的に対処すべき論点を示す手がかりとなる。逆に、プロセス指標とアウトカムが連動して改善している場合には、当初の接続仮説が妥当であったことの傍証となり、同種の業務への横展開の根拠として活用できる。

以上のように、補完手法の接続設計は、ログ中心の測定の仕組みを「プロセス・アウトプット層の観測装置」として明確に位置づけたうえで、その先にあるアウトカム層との間に仮説と検証の仕組みを設けるものである。これにより、本モデルは、ログで測れる範囲を最大限に活用しつつ、ログだけでは到達し得ない行政価値の把握にも開かれた評価体系として機能することが可能となる。

4.6 小括：デジタル時代の行政評価はデジタルで測るモデルへ

第一に、本章では、行政 DX の成果を把握する前提として、観測データと非観測データを区別する必要があることを示した。ログから直接把握できる処理時間や差戻し回数等は、継続観測の土台となる一方で、窓口での説明、電話対応、個別判断、住民の納得感、職員の心理的負担などは、ログのみでは十分に把握できない。したがって、本モデルはログを万能な測定手段として位置づけるものではなく、ログで安定的に測れる範囲を基礎としつつ、ログで捉えきれない価値については、実測、ヒアリング、アンケート等によって補完することを前提としている。

第二に、本章では、業務ログに基づく処理時間・滞留状況の把握とボトルネック特定の間考え方を整理した。受付、審査開始、差戻し、決裁、完了といったイベントを案件単位で連結することで、業務全体のリードタイムや工程別の待機時間を把握できる。また、差戻し回数、再処理頻度、引継ぎ回数、操作回数等を組み合わせることで、処理遅延の主因がどの工程にあるのか、あるいは前段階の入力・確認不足や部門間連携に起因しているのかを可視化できる。このような工程分析は、経験的な判断に依存しがちであった業務改善の対象を、データに基づいて優先順位付けするための土台となる。

以上のように、本章で提示したモデルは、行政 DX の効果検証を、追加的な調査負担に依存するものから、日常業務の中で生成されるデータを活用した継続的な評価の土台へと転換するものである。ただし、ログ中心の測定には、ログから見えない情報、プライバシー保護、監視リスク、分析・データ管理コストといった課題が残る。次章では、本研究全体の知見を踏まえ、自治体および関係機関が実務上取り組むべき提言を整理するとともに

に、本モデルの限界と今後の課題を明らかにする。

加えて、本モデルの射程に関して一点補足する。本モデルが直接観測するのはプロセス指標およびアウトプット指標であり、住民の満足度や職員の心理的負担といったアウトカム層を直接測定するものではない。本章では 4.5 節において、ログ観測とアウトカム評価を接続するための補完手法の設計枠組みを示した。すなわち、プロセス指標の改善がアウトカムの変化に到達しているかを、接続仮説の明示と年次サイクルへの補完手法の組み込みによって検証する仕組みである。本モデルは、ログで安定的に把握可能なプロセス・アウトプット層の継続観測に特化した測定の仕組みとして位置づけられるとともに、補完手法との接続を通じて、アウトカム評価にも開かれた構造を備えている。

5.まとめ

本章では、本研究全体の知見を整理し、行政 DX の実効性をどのように測定し、行政運営へ生かしていくべきかについて提言を述べる。ログを活用したデジタル測定の可能性とその限界を示すとともに、プライバシー保護や監視リスクに配慮した運用の重要性を示す。また、今後の検討課題や研究の発展方向を提示し、行政 DX の成果を持続的に可視化するための次のステップを示す。本章は、本研究の総括として、行政評価の新たな方向性を提案する役割を担う。

5.1 本研究の提言

本研究は、行政 DX の効果検証が指標の未整備・データ不足・前後比較や自治体間比較の困難といった構造的課題を抱えてきたことを確認したうえで、業務の中で自然に生成されるログ等のデータを活用し、業務と測定を一体化する低負荷・継続的なデジタル測定による効果検証モデルを提案してきた。本節では、ログを継続観測の土台としつつ、他の手法による補完を組み合わせることを前提として、自治体および関係機関が実務において取り組むべき事項を、以下の五点の提言として整理する。

提言 1：業務と測定を一体化し、追加負担の少ないデジタル測定へ転換する

従来の効果検証は、ヒアリングや実測など、本来業務とは別の追加的な作業に依存しがちであり、現場の負担と測定の継続性との間にトレードオフを抱えていた。本研究は、業務遂行の過程で生成されるログを測定の仕組みに据えることで、追加的な計測作業を抑えながら、案件単位・工程単位の処理実態を継続的に把握できることを示した。ここで測定の仕組みとなるのは、デジタル化された業務がその遂行過程で自然に生み出すデータ（本報告ではこれを「プロセス副産物データ」と呼ぶ）であり、システムログはその最も手の届きやすい一例である。重要なのは、測定のために現場へ追加の観測作業を課すのではなく、業務の副産物として取得できる範囲を見極めることであり、「追加負担ゼロないし極小で測れる範囲はどこまでか」を測定設計の境界線として据えることである。したがって、効果検証は事後的・一時的な調査としてではなく、業務プロセスそのものに測定機能を内在化させる設計思想のもとで構築されるべきである。新たな施策やシステムの導入にあたっては、企画段階から「何を、どのログで、どの粒度で観測するか」を組み込むことを標準的な考え方とすべきである。

提言 2：継続観測データを比較可能な共通の土台として整備し、改善に活用する

効果検証を個別自治体内の改善にとどめず、自治体横断の学習の土台へと発展させるためには、観測対象・定義・記録粒度の標準化が不可欠である。ただし、法令根拠が共通す

る業務であっても運用には差異があるため、厳密な完全統一を目指すのではなく、観測結果を相対比較可能な水準へ引き上げるための最低限の定義・記録仕様を整備することが現実的である。そのうえで、業務時間・件数・コスト・差戻し等の共通指標を蓄積し、成功事例のみならず期待した効果が得られなかった事例も共有する横断データ基盤を整備することで、各自治体は自らの水準を相対化し、試行錯誤の重複を避けながら改善の優先順位を判断できるようになる。

提言 3：「監視にならない測り方」をガードレールとして制度的に担保する

ログの活用は、設計意図が業務改善であっても、現場からは個々の職員に対する行動観察として受け止められ、測定への不信や萎縮を招く可能性がある。効果検証を持続的に運用するためには、ログ活用の目的を業務改善とサービス品質向上に限定し、個人の能力評価や監督目的に転用しないことを原則として明文化する必要がある。具体的には、取得目的の限定、最小限収集、匿名化・仮名化、アクセス統制、監査記録といった規律を整備するとともに、データや AI による分析はあくまで判断支援であり、最終的な判断と責任は職員が担うという原則（いわゆる Human-in-the-Loop）を堅持することを提言する。

提言 4：ログで測れない価値を補完し、創出時間の再配分・価値還元まで評価する

本調査における定量指標は、行政 DX の価値を数字だけで評価するためのものではなく、業務の状態変化を継続的に把握し、比較・分析・改善を支える共通のものさしとして位置づける。一方で、住民の満足度や幸福度といった主観的価値、職員の心理的負担や健康状態、窓口での説明や個別判断など、ログには現れにくい領域を取りこぼす。これらについては、実測・ヒアリング・アンケート等を補完的に組み合わせ、測定の射程を「効率性」から「負担・健康・価値」へと広げる必要がある。その際、補完手法は「別途実施する」という宣言にとどめず、4.5 節で示したとおり、ログ指標との接続仮説の明示、年次サイクルへの実施タイミングの組み込み、ログ指標との突合による乖離の検出という設計のもとで運用されるべきである。さらに、効果検証の目的は時間や費用の削減それ自体ではなく、創出された時間が住民サービスの質の向上や、相談・確認・予防・企画といったコア業務へ再配分・再投資される点にある。削減を終点とせず、価値の還元までを評価の射程に含めることを提言する。具体的には、削減された業務時間が、単に余剰時間として処理されるのではなく、住民対応、相談支援、確認・審査の質向上、企画・改善業務等へどのように再配分されたのかを把握する必要がある。創出時間の価値還元を評価対象に含めることで、行政 DX の成果を時間短縮にとどめず、住民サービスの質や職員のコア業務への集中に接続して捉えることができる。

提言 5：小さく始め、年次サイクルを通じて制度・計画に組み込む

効果検証モデルは、最初から全庁横断・自治体横断に一律導入するのではなく、限定的な

対象業務における PoC から着手し、既存システムの記録や推計による暫定的なベースラインを活用しながら、低コストに観測を立ち上げることが現実的である。そのうえで、観測 → 分析 → 改善 → 再観測の年次サイクルを定着させ、対象工程の拡張と指標の洗練を段階的に進めるべきである。最終的には、調達仕様・ログ設計・定義書管理・引継ぎ資産・年次固定費化といった制度・運用の仕組みに組み込み、観測と改善のサイクルを事業計画や基本計画に明確に位置づけることで、担当者やシステムが変わっても継続する評価の土台として制度化することを提言する。

以上の五点は相互に独立したものではなく、ログを前提とした測定設計（提言 1）を、最低標準化と横断基盤（提言 2）、監視にしないためのガードレール（提言 3）、ログで測れない価値の補完（提言 4）によって支え、段階的な導入と制度化（提言 5）によって持続させるという、一体の枠組みを構成している。

5.2 本研究の限界と課題

本研究が提案するログを基点とした効果検証モデルは、業務遂行の過程で自動生成されるデータを継続的・低負荷に活用できる点に強みを持つ。しかし、その一方で、ログという測定の仕組みそのものに由来する限界や、運用上留意すべき課題も存在する。本節では、本モデルの限界と課題を、ログから見えない情報（5.2.1）、監視にならない測り方の導入（5.2.2）、プライバシー保護の観点（5.2.3）、分析・データ保存コスト（5.2.4）の四点から整理する。

5.2.1 ログから見えない情報

業務ログを基盤とする効果検証は、処理時間、待機時間、処理件数、差戻し回数といった、業務遂行の過程で痕跡として残る定量的情報の把握に優れている。しかし、行政サービスの価値や職員の業務実態には、ログには現れにくい、あるいは原理的に記録されない領域が存在する。

第一に、システム外で行われる業務である。窓口での口頭による説明、電話やメールを通じた個別の応対、案件ごとの裁量的判断などは、その多くがシステムの操作履歴として残らない。また、紙媒体を前提とした業務や、デジタル化以前の工程については、そもそもログとして取得する対象が存在しない。職員の移動を伴う現地対応や、車両の運行・給油といった現場業務も、業務システムのログだけでは捕捉できない。

第二に、住民にとっての主観的な価値である。手続に要する時間や回数といった客観指標は改善の有無を示す手がかりとなるが、住民が実際に感じる利便性、安心感、満足度、さらには地域全体の幸福度といった価値は、ログのみでは測定できない。住民の総幸福度の

ような主観指標を行政評価に取り入れようとする試みも見られるが、こうした価値はアンケートやヒアリングを通じてはじめて把握される。たとえば、荒川区では、荒川区民総幸福度（GAH）として46の指標を構築し、毎年のアンケート調査を通じて5段階評価で定量化している。その結果は、特定の施策との因果関係を直接示すものとしてではなく、住民ニーズの把握や政策立案における補助的な証拠として活用されている。実際に、防災分野における実感が低い層が特定されたことを契機として、地域の防災イベントの企画につながった事例が報告されている。このように、主観指標はログベースの効率性評価とは異なる射程を持ち、政策全体の方向性を示す補完的な役割を果たし得る。

第三に、職員の心理的・身体的な負担である。業務量や処理時間の削減がログ上で確認できたとしても、その過程で職員が感じる心理的ストレスや健康面への影響は、操作履歴には反映されない。研修や業務体験を通じた負荷の把握、あるいはストレスや生理指標の測定といった、ログの外側にある手法によってはじめて捉えられる側面がある。

したがって、本モデルは「ログで測れるものから固める」ことを基本としつつ、ログで把握できないこれらの領域については、実測・ヒアリング・アンケート・主観指標・生理指標等を補完的に組み合わせ、測定の射程を「効率性」から「住民価値・負担・健康」へと広げる必要がある。ログ中心の測定は万能ではなく、補完的手法と役割分担を明確にしたうえで運用されるべきである。

なお、本研究では、こうしたログから見えない情報を単に限界として指摘するにとどめず、4.5節においてログ観測と補完手法の接続設計を具体的に示した。すなわち、ログで把握可能なプロセス指標と、補完手法によって検証すべきアウトカムとの間に接続仮説を明示し、年次サイクルの中で両者を突合する枠組みを設けることで、ログの限界を構造的に補う設計としている。

5.2.2 監視にならない測り方の導入

ログを用いた効果検証は、その設計意図が業務改善にあったとしても、現場の職員からは個々の行動を観察・評価する仕組みとして受け止められる可能性がある。とりわけ、端末操作ログや画面遷移履歴のように個人の作業過程に近い情報を扱う場合、測定への不信や萎縮を招き、記録・分析の運用そのものが組織内で受容されにくくなる。効果検証を持続的に機能させるためには、「監視にならない測り方」を制度的に担保することが不可欠である。

第一の原則は、目的の限定である。ログ活用の目的を業務改善とサービス品質の向上に限定し、個々の職員の能力評価や勤務監督の目的に転用しないことを、あらかじめ明文化する必要がある。測定の対象は、処理工程や機能利用状況といった比較的客観性の高い領域から段階的に開始し、個人を評価するための指標として用いないことを徹底する。

第二の原則は、人による判断の留保である。データや AI による分析は、あくまで職員の判断を支援するための材料であり、最終的な判断と責任は職員が担う（Human-in-the-Loop）。分析結果が機械的に評価や処遇へ直結する運用を避けることで、測定が現場を萎縮させる装置になることを防ぐ。

第三の原則は、受容性を高める運用設計である。新たな監視的仕組みを構築するのではなく、既存の監査ログなど、すでに業務上正当な目的で取得・運用されている記録を起点とすることで、現場の心理的抵抗を抑えることができる。あわせて、データ取得についての同意、閲覧権限の限定、分析結果の利用範囲の明示、参加を望まない職員への配慮といった運用上の手当てを講じることが望ましい。さらに、成功事例だけでなく期待した効果が得られなかった事例も共有し、失敗を責めない組織文化を醸成することが、測定への信頼を支える。

もっとも、受容性の確保は、制度設計の段階で完了するものではなく、運用の過程で段階的に醸成されるものである。実務においては、業務データの可視化が導入される際に、処理時間の公開が個々の職員の能力評価に転用されるのではないかという心理的懸念が現場に生じることは珍しくない。この懸念に対しては、データの利用目的が個人評価ではなく業務全体の改善にあることを組織として明確に宣言し、分析結果の開示範囲や利用場面をあらかじめ限定することが第一歩となる。しかし、宣言だけでは十分ではない。重要なのは、宣言した目的が実際の運用の中で一貫して守られていることを、現場が経験的に確認できるプロセスを経ることである。データが改善のために使われ、個人の責任追及には使われないという実績が積み重なることで、職員の間でデータを見る文化が徐々に浸透し、やがては現場側から改善提案が生まれるようになる。すなわち、受容性とは「懸念を抑制する」ことではなく、「データ活用の目的と実績の一致を通じて信頼を構築する」ことであり、その設計には一定の時間と運用上の一貫性が求められる。

これらは、本研究が 4.3.3 で示した「小さく始めて継続するための運用ガイドライン」におけるガードレールを、限界・課題の観点から改めて整理したものである。取得目的の限定、個人評価への不転用、最小限収集、匿名化・仮名化、アクセス統制といった具体的規律のうち、データ保護に関わる側面は次の 5.2.3 で述べる。

5.2.3 プライバシー保護の観点

ログを基盤とする効果検証は、業務に関する記録だけでなく、場合によっては職員や住民の行動に関する機微な情報を含みうる。測定を正当かつ持続的に運用するためには、プライバシー保護のための規律を制度として整えることが前提となる。本研究で想定するログ活用は、個人情報保護法その他関係法令及び各自治体の個人情報保護制度等との整合を前提として実施されるものである。

第一に、最小限収集の原則である。効果検証に必要な範囲を超えてデータを収集せず、目的に照らして必要なログ項目・粒度に限定する。第二に、匿名化・仮名化である。分析

にあたっては、個人を直接特定する情報を排し、必要に応じて仮名化を施すことで、個人の評価や特定に利用されないようにする。第三に、アクセス統制と閲覧権限の限定である。誰がどのデータにアクセスできるのかをあらかじめ明確に定め、権限を持つ者以外が機微な情報に接することのないよう統制する。第四に、監査記録の保持である。データの取得・利用・分析の過程を記録し、運用が定めた目的・範囲の内にあることを事後的に検証可能な状態に保つ。

特に、ストレスや生理指標等の健康に関わる機微なデータを活用する場合には、本人に対する十分な説明と同意、利用目的の明確化及び適切な安全管理措置を前提とする。また、当該データが職員の不利益な取扱いに転用されないよう、閲覧権限及び利用範囲を厳格に限定する必要がある。また、業務上取得したログを当初の目的を超えて二次的に利用する場合には、対象となる職員や関係者との間で改めて合意を形成することが求められる。これらのデータ保護規律は、5.2.2 で述べた「監視にならない測り方」と一体のものであり、両者を併せて整備することによって、効果検証が現場に受容され、継続的に運用される基盤が確保される。

5.2.4 分析・データ貯蔵コスト

加えて、本研究が前提とするログ活用やプロセスマイニングの実装可能性は、技術環境や導入時点の市場条件にも左右される。特に初期段階においては、利用可能な分析ツールが限定され、導入・運用コストが高額となる場合があったため、すべての自治体が同一条件で実装可能であったわけではない。したがって、効果検証モデルの実務適用にあたっては、技術的可能性のみならず、費用対効果、運用体制、既存システムとの接続可能性を踏まえた段階的導入を前提とする必要がある。

5.3 展望

本研究が提示したログを基点とした効果検証モデルは、直ちに全自治体・全業務へ一律導入されるべき完成形としてではなく、段階的に実装・拡張されるべき発展的な枠組みとして理解する必要がある。今後の展望は、大きく短期・中期・長期の三段階で整理できる。これは、第 4.3 節で示した導入ステップ、すなわち最低要件の整備、年次サイクルの運用、小さく始めて継続する段階的拡張と対応している。

短期的には、既存システムにすでに蓄積されているログや監査記録等を活用し、PoC として限定的な対象業務から観測・分析を開始することが現実的である。この段階では、case/activity/timestamp/resource の最小スキーマを満たす範囲で対象を絞り、処理時間、待機時間、差戻し等の基本指標を把握し、可視化と改善の接続可能性を確認することが主眼となる。すなわち、まずは「ログで測れるものから固める」ことで、測定の持続可能性と実務上の有用性を具体的に示す段階である。なお、改修前のデータが存在せず厳密な前

後比較が難しい場合であっても、改修前の状態を推計で補った暫定的なベースラインを設定することで、効果把握の出発点を確保することができる。

中期的には、対象業務や対象部門を拡張し、観測 → 分析 → 改善 → 再観測の年次サイクルを定着させることが課題となる。この段階では、単一業務内の改善確認にとどまらず、部門横断的な工程分析や、共通指標・共通語彙に基づく自治体間比較の可能性が広がる。横断データ基盤の整備が進めば、自団体の結果を他団体と相対化し、改善余地をより精緻に把握できるだけでなく、成功・失敗の双方の知見を相互に参照する学習循環も成立し得る。

長期的には、こうした取組を個別の PoC や先進団体の実践にとどめず、調達仕様、ログ設計、定義書管理、引継ぎ資産、年次固定費化といった制度・運用の仕組みに組み込み、行政 DX の評価基盤として制度化していくことが求められる。その際には、観測・改善のサイクルを単年度限りの取組として終わらせず、事業計画や基本計画に明確に位置づけ、制度的・予算的な裏付けを確保することが、継続性を担保するうえで重要となる。すなわち、ログ出力やエクスポート、ID 連結、スキーマ版管理、匿名化等の要件を初期設計段階から内在化し、年度や担当者が変わっても継続観測が維持される状態を目指す必要がある。その先には、自治体間で比較可能な共通指標のもとで、行政 DX の成果を横断的かつ継続的に可視化し、改善知見を広く横展開していく基盤の形成が展望される。

もっとも、このような制度化は、技術的可能性のみで実現するものではない。とりわけ、住民満足度や幸福度といった主観的価値、あるいは職員の心理的負担や健康状態など、ログには現れにくい側面については、実測・ヒアリング・アンケート等を補完的に組み合わせる必要がある。こうしたログで把握できない価値の補完に加え、監視への懸念への対応、プライバシー保護、分析・保存コストの確保といった論点を引き続き調整しながら、測定と改善を両立する運用原則を洗練させていくことが不可欠である。その意味で、本研究が提案するモデルは到達点ではなく、行政 DX を「デジタルで測り、改善に継続する」ための実装可能な出発点として位置づけられる。

ヒアリング先自治体

本調査研究にあたり、行政 DX の効果を検証することの難しさを把握するために様々な自治体へのヒアリングを実施した。この場を借りて、快くヒアリングを受けていただいた自治体の方々に深く感謝申し上げます。

1. 群馬県前橋市 未来創造部 情報政策課
2. 東京都板橋区 政策経営部 IT 推進課
3. 東京都練馬区 区民部 収納課
4. 東京都多摩市 企画政策部 情報政策課

5. 東京都荒川区 政策企画部 政策企画課
6. 同上 政策企画部 デジタル推進課
7. 公益財団法人 荒川区自治総合研究所
8. 東京都江戸川区 児童相談所 援助課
9. 静岡県裾野市 総務部 情報統括課

注

本報告では、行政 DX の取組によって生じた変化を把握し、その妥当性を確認したうえで業務改善や政策立案に活用する一連の営みを「効果検証」と総称する。このうち「指標」は何を把握するかを定めるものであり、「測定」は当該指標を観測・記録・算出する方法を指す。また「評価」は、得られた結果を政策目的や行政価値に照らして解釈し、意味づけることを指す。なお、総称としての「効果検証」に含まれる「検証」は、厳密な因果推論のみを意味するものではなく、指標設計・測定・評価を通じて行政 DX の成果を適切な形で把握することを含む広義の概念として用いる。類語である「効果測定」「効果評価」は、文脈上特に必要な場合を除き、原則として「効果検証」に統一する。

参考文献

- さいたま市. “固定資産税評価業務に AI を活用します！” 2020.
- つくば市／日本電気株式会社. “市民窓口業務改革タブレットで転入届作成共同研究実績報告書.” 2020 年.
- 塩尻市. “保育業務改革プロジェクト（RPA・AI 等の活用による効率的な業務運営の実現）報告書.” 総務省. 2019 年 2 月. https://www.soumu.go.jp/main_content/000611121.pdf [アクセス日: 2026 年 6 月 16 日].
- 江戸川区. VR を活用した児童相談所職員ロールプレイ研修. 2025. <https://www.digitalservice.metro.tokyo.lg.jp/documents/d/digitalservice/entrysheet-dx2025#page=45> [アクセス日: 2026 年 6 月 16 日].
- 荒川区自治総合研究所. あたたかい地域社会を築くための指標（荒川区民総幸福度：GAH）. 2015. https://rilac.or.jp/?page_id=319 [アクセス日: 2026 年 6 月 16 日].
- 静岡県裾野市. “自治体フロントヤード改革モデルプロジェクト.” 総務省. 2025 年 3 月. https://www.soumu.go.jp/main_content/001007481.pdf [アクセス日: 2026 年 6 月 16 日].

泉大津市. “泉大津市業務改革推進プロジェクト 報告書.” 総務省. 2019 年 2 月. https://www.soumu.go.jp/main_content/000611111.pdf [アクセス日: 2026 年 6 月 16 日].

船橋市. “業務改革モデルプロジェクト（窓口業務プロセス改革事業）報告書.” 2019.

前橋市グループ. “報告書.” 総務省. 日付不明. https://www.soumu.go.jp/main_content/000678570.pdf [アクセス日: 2026 年 6 月 16 日].

足立区. “申請手続き等処理業務における RPA を活用した業務改革モデルの検証報告書.” 総務省. 2019 年 2 月. https://www.soumu.go.jp/main_content/000611324.pdf [アクセス日: 2026 年 6 月 16 日].

町田市. “自治体間ベンチマーキング ～業務の見える化とその先にあるもの.” 日付不明.

東京都. “東京都住民基本台帳人口移動報告.” 日付不明.

東京都デジタルサービス局. “東京都行政手続きデジタル化 QOS 向上ガイドライン.” 2024.

板橋区. “板橋区統合型地理情報システム導入・活用計画.” 2012.

練馬区. “～住民税等の未納対策強化と生活再建の支援に向け～ 練馬区と富士通 Japan で共同開発した"未納対策お支援する AI"を全国で初めて開始します.” 練馬区. 2024 年 3 月 27 日. <https://www.city.nerima.tokyo.jp/kusei/koho/hodo/r6/r603/20240327.files/20240327.pdf> [アクセス日: 2026 年 6 月 26 日].

和歌山県橋本市. “R P A ・ A I O C R による窓口業務改革及び県をまたぐ広域展開モデル推進事業実施結果報告書.” 総務省. 2019 年 2 月. https://www.soumu.go.jp/main_content/000611435.pdf [アクセス日: 2026 年 6 月 16 日].

用語集

行政 DX

行政分野において、デジタル技術を活用し、行政サービスや業務プロセスのあり方を見直すことで、住民・事業者の利便性向上、職員の負担軽減、組織運営の改善を図る取組。

効果検証

行政 DX の取組によって生じた変化を把握し、その妥当性を確認したうえで、業務改善や政策立案に活用する一連の営み。

デジタル測定

デジタル技術によって生成・記録されるデータを活用し、業務やサービスの実行状況を継続的に把握すること。本報告では、主としてログデータを用い、追加的な記録負担を抑えながら、処理時間、待機時間、差戻し回数、処理件数等を低負荷かつ継続的に観測する考え方を指す

ベースライン

ベースライン（baseline）とは、施策・介入の効果を測定する際の比較基準となる状態を指す。本研究では二つの意味で用いる。第一に、事前ベースラインとして、施策導入前の業務量・処理時間・件数等を定量的に記録した基準値を指す。第二に、反実仮想ベースライン（counterfactual baseline）として、施策が実施されなかった場合に想定される推移を指し、外的要因（制度変更・需要変動・人員増減等）の影響を分離するための仮想的な比較線である。効果は、事後の観測値とこれらベースラインとの差分として推定される。デジタル測定によって継続的にログを取得することは、両ベースラインの構築精度を高める前提となる。

観測可能領域

観測可能領域とは、業務の過程や成果のうち、外部から把握可能な形で現れている部分を指す。ログデータや計測値として記録される行動や結果がこれに含まれる。一方で、判断の意図や主観的評価など、外部から直接把握できない要素は含まれない。

非観測領域

非観測領域とは、業務の過程や成果のうち、外部から直接把握できない部分を指す。判断の意図、納得感、心理的負担など、行動として外部に現れない要素が含まれる。これらはログデータ等による継続的な観測が困難であり、アンケートやヒアリング等の補完的手法によって把握される。

因果推論

因果推論とは、観測された変化が特定の施策や介入によって生じたものかを識別するための分析を指す。単なる前後比較ではなく、施策が存在しなかった場合に想定される状態（反実仮想）と比較することで、外的要因の影響を分離し、施策の効果を推定する。本研究では、ログデータ等による継続的な観測を前提としつつ、必要に応じて差の差法（DID）等の手法を適用することで、効果の解釈可能性を高めるものとして位置づける。

反実仮想

反実仮想とは、施策や介入が実施されなかった場合に想定される状態を指す。実際に観測された結果と比較することで、施策による変化を識別するための基準となる。現実には直接観測することができないため、比較群や過去データ等を用いて推定される。

EBPM（Evidence-Based Policy Making）

政策立案や行政運営において、客観的な証拠（エビデンス）に基づいて意思決定を行う枠組み。行政 DX の効果検証や評価設計にも応用されるが、政策分野全般で用いられる普遍

的な概念。

プロセスマイニング

イベントログを用いて、実際の業務プロセスを自動的に発見・分析・評価する手法。発見（discovery）、適合性チェック（conformance checking）、パフォーマンス分析などの技術を通じて、業務の可視化や改善、予測・自動化につなげる。van der Aalst により体系化された

指標の定義書

観測や分析に用いる指標の意味、計算方法、単位、集計ルールなどをまとめた一覧表や文書。