Logiki instytucjonalne i bariery zarządzania wpływem w świetle koncepcji Matta Stevensona-Dodda

AUTOR

Fundacja Dobre Państwo

STRESZCZENIE / ABSTRACT

Tekst analizuje koncepcję Matta Stevensona-Dodda, podkreślając, że systemy pomiaru wpływu (Understand-Measure-Maximise) muszą być osadzone w specyficznej logice instytucjonalnej danego sektora. Autor wskazuje na różnice między biznesem, administracją a NGO, gdzie role płatnika i beneficjenta są rozdzielone. Kluczowym postulatem jest integracja wpływu z rdzeniem modelu biznesowego zamiast ograniczania go do działań PR czy CSR. Artykuł wprowadza istotne rozróżnienie pomiędzy sprawozdawczością ESG/ESRS (materialność) a rygorystyczną ewaluacją przyczynową (causal impact evaluation). Wskazuje również na konflikt między materialnością finansową a społeczną, co determinuje sposób raportowania i realnego zarządzania zmianą w organizacjach.

The text analyzes Matt Stevenson-Dodd's concept, emphasizing that impact measurement systems (Understand-Measure-Maximise) must be embedded in the specific institutional logic of a given sector. The author points out differences between business, administration, and NGOs, where the roles of payer and beneficiary are separated. The key postulate is the integration of impact into the core of the business model instead of limiting it to PR or CSR activities. The article introduces an important distinction between ESG/ESRS reporting (materiality) and rigorous causal impact evaluation. It also indicates a conflict between financial and social materiality, which determines the method of reporting and real change management within organizations.

Artykuł analizuje wyzwania związane z implementacją modelu zarządzania wpływem społecznym (Understand-Measure-Maximise) w różnych sektorach instytucjonalnych,

ze szczególnym uwzględnieniem specyfiki polskiego systemu. Autor stawia tezę, że skuteczny pomiar wpływu nie może być uniwersalnym szablonem, lecz musi wynikać z logiki danej organizacji – od biznesu i impact investingu, przez administrację publiczną, aż po sektor NGO. Głównym problemem nie jest brak narzędzi analitycznych czy danych (które w Polsce są często nadmiarowe), lecz zjawisko 'decouplingu' – rozdzielenia formalnego raportowania od rzeczywistych procesów decyzyjnych. Tekst argumentuje, że przejście od kultury 'wykonania zadania' (spending completion) do kultury 'osiągnięcia zmiany' (mission completion) wymaga nie tyle nowych wskaźników KPI, co stworzenia bezpiecznej infrastruktury do uczenia się na błędach i odwagi w korygowaniu alokacji zasobów na podstawie rzeczywistych rezultatów, a nie tylko zgodności proceduralnej.

SŁOWA KLUCZOWE

logiki instytucjonalne | zarządzanie wpływem | Matt Stevenson-Dodd | Theory of Change | causal impact evaluation | impact investing | ESRS reporting | performance budgeting | resource dependence theory | mission drift | additionality inwestora | accountability for learning | audit Theory of Change | social materiality | financial materiality

Konieczność osadzenia modelu wpływu w logice instytucjonalnej

Jeden język wpływu, różne logiki instytucjonalne: biznes, państwo, organizacje społeczne, inwestorzy, fundatorzy i sport

Wstępna analiza Fundacji Dobre Państwo traktowała „organizację” jako kategorię względnie jednolitą. Choć na etapie rekonstrukcji metody Stevenson-Dodda było to rozwiązanie użyteczne, dalsza część wywodu wymaga korekty.

Przedsiębiorstwo giełdowe, urząd centralny, gmina, fundacja charytatywna, przedsiębiorstwo społeczne, fundusz impact investing czy klub sportowy nie są jedynie różnymi rozmiarami tego samego typu podmiotu. Funkcjonują one w odmiennych systemach własności, odpowiedzialności, finansowania i prawa, a także w różnych ramach legitymizacji oraz dystrybucji władzy. W rezultacie

ten sam termin – **impact** – pełni w każdym z nich nieco inną funkcję. Autor książki słusznie dostrzega tę sektorową heterogeniczność, analizując osobno biznes, administrację, organizacje charytatywne i przedsiębiorstwa społeczne, inwestycje społeczne, fundatorów oraz sport.

Kluczowa konkluzja nie sprowadza się jednak do stworzenia sześciu różnych podręczników pomiaru. Chodzi raczej o zrozumienie, że architektura Understand-Measure-Maximise musi zostać osadzona w logice instytucjonalnej konkretnego sektora. Nie można mechanicznie przenosić systemu właściwego inwestorowi do organizacji obywatelskiej, modelu przedsiębiorstwa do urzędu czy metod zarządzania funduszem venture capital do pomocy społecznej.

Różnice sektorowe najłatwiej uchwycić poprzez pięć pytań (bez konieczności zamykania ich w formalnej tabeli). Należy ustalić: kto definiuje cel organizacji, kto jest odbiorcą jej działań, kto za nie płaci, wobec kogo organizacja ponosi odpowiedzialność oraz kto ma prawo zmieniać alokację zasobów. W przedsiębiorstwie klient, płatnik i właściciel mogą być różnymi osobami. W administracji płatnikiem jest podatnik, odbiorcą obywatel, a mandat do działania wynika z prawa. W NGO beneficjent często nie jest ani donatorem, ani formalnym właścicielem. Z kolei w impact investing inwestor dostarcza kapitał, lecz deklarowanym adresatem wpływu pozostaje osoba trzecia. To właśnie te rozszczepienia decydują o tym, jakie informacje są produkowane, komu służą i jakie zachowania wywołuje sam system pomiarowy.

Integracja wpływu z rdzeniem modelu biznesowego

Autor przeciwstawia tradycyjny CSR bardziej strategicznemu podejściu, w którym wpływ społeczny zostaje zintegrowany z rdzeniem modelu biznesowego, alokacją kapitału i decyzjami zarządu. Krytykuje sytuację, w której działania społeczne są oddelegowane do PR lub raportów ESG, podczas gdy podstawowa działalność przedsiębiorstwa pozostaje wyłączona z analizy wpływu. Ta intuicja wymaga jednak ważnego rozszerzenia: w firmie wpływ nie może być analizowany wyłącznie przez pryzmat programów dobroczynnych. Znacznie większe skutki społeczne mogą bowiem wynikać z tego, jaki jest **produkt, warunki pracy, wynagrodzenia, łańcuch dostaw, polityka podatkową, technologia, sposób wykorzystywania danych, środowisko oraz siła rynkowa przedsiębiorstwa**. Firma może zatem prowadzić znakomity program społeczny i jednocześnie generować negatywne externalities wielokrotnie przewyższające pozytywne rezultaty filantropii. Z punktu widzenia Theory of Change ograniczenie impact assessment do fundacji korporacyjnej byłoby błędem w wyborze jednostki analizy. Należy objąć tym samym aparatem zarówno skutki zamierzone, jak i uboczne, pozytywne i negatywne, które powstają nie tylko "obok biznesu", ale przez biznes.

Właśnie dlatego współczesna europejska sprawozdawczość zrównoważonego rozwoju wykracza poza dawny model CSR. W lipcu 2026 r. Komisja Europejska przyjęła zrewidowane ESRS, które – mimo uproszczenia obowiązków po pakiecie Omnibus I – mają dostarczać informacji o oddziaływaniu przedsiębiorstw na ludzi i środowisko oraz o ryzykach związanych ze zrównoważonym rozwojem. Nie wolno jednak utożsamiać tego języka regulacyjnego z ewaluacją przyczynowego wpływu programu społecznego. Ustalenie, że dana kwestia jest materialnym oddziaływaniem przedsiębiorstwa, nie jest tożsame z oszacowaniem kontrfaktycznego efektu konkretnej interwencji. To rozróżnienie jest kluczowe: **ESG reporting, impact materiality i causal impact evaluation nie są synonimami.** Pierwsze dwa pojęcia pomagają opisać, jakie skutki i ryzyka są istotne z punktu widzenia przedsiębiorstwa i jego otoczenia.

Trzecie odpowiada na bardziej rygorystyczne pytanie: jaka zmiana zaszła dzięki określonemu działaniu w porównaniu z tym, co wydarzyłoby się bez niego. Przedsiębiorstwo może więc posiadać bardzo rozwiniętą sprawozdawczość ESRS i nadal mieć słabą wiedzę o przyczynowej skuteczności konkretnego programu społecznego. W biznesie pojawia się także szczególny konflikt pomiędzy **financial materiality i social materiality**. Wpływ, który jest ogromny dla niewielkiej społeczności, może być finansowo niematerialny dla globalnej korporacji. Jeżeli wszystkie problemy społeczne zostaną filtrowane wyłącznie przez ich wpływ na wynik przedsiębiorstwa, interesariusze o małej sile ekonomicznej mogą zostać epistemicznie zepchnięci na margines. Z drugiej strony system wpływu całkowicie oderwany od ekonomiki firmy może pozostać dekoracją pozbawioną zdolności do rzeczywistej zmiany decyzji. Dlatego koncepcja "impact is strategy" jest w przedsiębiorstwie najbardziej wiarygodna wtedy, gdy oddziałuje na **capex, produkt, pricing, politykę zatrudnienia, procurement i konstrukcję łańcucha wartości**, a nie tylko na wielkość budżetu CSR. Dopiero wtedy dane o wpływie zaczynają konkurować o uwagę z ROI, EBITDA, ryzykiem i wzrostem.

Jednocześnie biznes posiada specyficzną przewagę nad częścią organizacji społecznych: dysponuje zwykle znacznie lepszą telemetrią operacyjną. Wie dokładnie, co, komu, kiedy i za ile sprzedał. Problem nie polega więc często na braku danych, lecz na braku połączenia wskaźników ekonomicznych z informacją o społecznych konsekwencjach działania. W tym sektorze impact management jest w dużej mierze problemem **integracji systemów informacyjnych i rozszerzenia funkcji celu**, a nie budowania infrastruktury danych od zera.

Autor przedstawia sektor publiczny jako obszar, w którym należałoby odejść od nadmiernego skupienia na procedurach ku ocenie rzeczywistego dobrostanu obywateli. Kierunek ten jest trafny, lecz wymaga szczególnej ostrożności.

Balans między legalnością a skutecznością w administracji

Proceduralność państwa nie jest jedynie przejawem biurokratycznego konserwatyzmu. Weberowska racjonalność formalna, legalność, przewidywalność decyzji, dokumentowanie podstaw działania oraz równe stosowanie prawa stanowią kluczowe zabezpieczenia przed arbitralnością. Państwo nie może zatem zostać przekształcone w organizację typu „Difference" poprzez proste zastąpienie procedur wskaźnikami outcome. Skuteczność nie legitymizuje działania nielegalnego, podobnie jak legalność nie dowodzi skuteczności; dojrzałe zarządzanie publiczne musi zapewniać oba te warunki jednocześnie. Współczesne praktyki OECD wskazują, że administracje rzeczywiście ewoluują w stronę performance budgeting. W 2023 roku 28 z 33 badanych państw OECD stosowało formę budżetowania zadaniowego lub wynikowego. Performance information służyło do corocznej alokacji środków w 71% tych państw, do decyzji menedżerskich w 57%, natomiast do wieloletniego planowania budżetowego już tylko w 32% przypadków.

Dane te obrazują różnicę między samym posiadaniem informacji o rezultatach a rzeczywistym włączeniem ich do architektury alokacyjnej. OECD podkreśla, że performance budgeting ma przesuwać punkt ciężkości z nakładów (inputs) na rezultaty, lecz nie jest on mechanicznym systemem cięcia funduszy według rankingu KPI. Informacja o wynikach powinna wspierać decyzje budżetowe i rozliczalność; w wielu państwach służy ona przede wszystkim do rozpoznawania przyczyn odchyleń, a nie do automatycznego sankcjonowania jednostek. Jest to istotne przeciwstawienie zbyt uproszczonej wizji Live Data jako algorytmu automatycznie przesuwającego środki. W administracji szczególne znaczenie ma również problem wielu principals. Organ publiczny odpowiada bowiem jednocześnie przed ustawodawcą, organami kontroli, sądami, ministrem, podatnikami, obywatelami oraz grupami bezpośrednio korzystającymi z usługi.

Definicje sukcesu tych podmiotów mogą pozostawać w napięciu. Skrócenie czasu postępowania może poprawić doświadczenie obywatela i obniżyć koszty, lecz nadmierna presja na szybkość grozi pogorszeniem jakości decyzji administracyjnych. Zwiększenie liczby świadczeń poprawia dostępność, ale podnosi ryzyko błędnych wypłat. Efektywna kontrola ogranicza nadużycia, lecz jednocześnie zwiększa koszt transakcyjny dla uczciwych użytkowników. Właśnie dlatego sektor publiczny potrzebuje balanced outcomes, a nie pojedynczego KPI. Mierniki skuteczności w administracji powinny być osadzone w ramach praworządności, równego traktowania, dostępności, jakości, procedural fairness oraz możliwości odwołania. Szczególnie interesującym obszarem są zamówienia publiczne. Autor książki proponuje projektowanie ich według zasady Function First,

czyli definiowanie społecznego rezultatu przed formą procedury. Podejście to znajduje częściowy odpowiednik w pracach Komisji Europejskiej dotyczących socially responsible public procurement. Opublikowany w 2025 roku unijny framework pomiaru społecznego wpływu zamówień, przetestowany na 75 procedurach z lat 2023–2024, obejmuje m.in. zatrudnienie, warunki pracy, dostępność oraz zrównoważony rozwój. Jednocześnie outcome-based procurement może generować własne problemy Goodharta.

Jeśli wykonawca otrzymuje premię za liczbę osób pozostających w pracy przez sześć miesięcy, pojawia się pokusa selekcji łatwiejszych uczestników. Dlatego kontrakt oparty na rezultacie nie jest automatycznie bardziej sprawiedliwy niż ten oparty na czynnościach. Jego jakość zależy od konstrukcji populacji, risk adjustment, zabezpieczeń przed zjawiskami creamingu i parkingu oraz od zdolności zamawiającego do odróżnienia wykonawcy słabego od takiego, który pracuje z grupą trudniejszą.

Zależność od fundatorów wymusza priorytetyzację raportowania nad wpływem

NGO i przedsiębiorstwa społeczne: pomiędzy misją, fundatorem a głosem beneficjenta

W organizacjach pozarządowych najbardziej charakterystyczną cechą jest rozdzielenie płatnika i beneficjenta. Osoba otrzymująca pomoc zazwyczaj nie finansuje usługi, natomiast fundator przekazuje środki, nie doświadczając bezpośrednio efektów interwencji. To strukturalne rozdzielenie tworzy warunki dla szczególnej formy principal-agent problem: organizacja stara się jednocześnie odpowiadać na realną potrzebę społeczną i na wymagania podmiotu finansującego. Autor książki trafnie wskazuje tutaj na ryzyko paraliżu analitycznego w sektorze społecznym – ograniczone zasoby zderzają się z rosnącą liczbą wymogów raportowych, a kwestia wpływu bywa delegowana osobom o niewielkim wpływie na strategię. Problem ten można wyjaśnić głębiej poprzez resource dependence theory. Organizacja zależna od zewnętrznego finansowania zaczyna częściowo dostosowywać swoją strukturę i język do oczekiwań dostawcy zasobu. Jeśli każdy fundator wymaga innego zestawu wskaźników, NGO produkuje wiele systemów raportowania nie dlatego, że uważa je za epistemicznie wartościowe, lecz dlatego, że każdy stanowi warunek dostępu do środków.

W skrajnej postaci prowadzi to do mission drift. Organizacja nie zmienia misji formalnie, jednak jej portfel projektów stopniowo przesuwa się w stronę tego, co jest możliwe do sfinansowania. W rezultacie Theory of Change może zostać odwrócona: zamiast poszukiwać zasobów dla strategii,

organizacja buduje strategię wokół dostępnych konkursów. Właśnie dlatego OECD i Komisja Europejska w aktualnych materiałach dotyczących social economy odchodzą od postulatu jednej uniwersalnej metodologii. Wspólny przewodnik z 2024 r. podkreśla, że pomiar powinien być dostosowany do misji i możliwości organizacji, stosować zasadę proporcjonalności, uwzględniać interesariuszy oraz służyć nie tylko accountability, ale przede wszystkim uczeniu się i poprawie działania. OECD wskazuje również, że praktyki impact measurement były w dużej mierze kształtowane przez potrzeby fundatorów i biznesu, co nie zawsze odpowiada specyfice social economy. Jest to istotne potwierdzenie i pogłębienie jednej z intuicji Stevensona-Dodda: proporcjonalność pomiaru jest częścią skuteczności, ponieważ koszt systemu ewaluacyjnego również obciąża misję.

Sytuację komplikuje specyfika przedsiębiorstwa społecznego, które funkcjonuje na skrzyżowaniu dwóch logik instytucjonalnych: rynkowej i misyjnej. Z jednej strony potrzebuje przychodów, klienta, dyscypliny kosztowej i zdolności inwestycyjnej, z drugiej zaś definiuje sukces poprzez rezultat społeczny. Te dwie logiki mogą współdziałać, ale mogą również wejść w konflikt – najbardziej rentowni klienci mogą nie być tymi, którym pomoc jest najbardziej potrzebna.

Pomiar wpływu jako ochrona misji i architektura bodźców

Skalowalny produkt może generować mniejszą głębokość zmiany. Inwestor często preferuje wzrost, podczas gdy misja wymaga intensywnej i kosztownej pracy z niewielką grupą osób. Dlatego w przedsiębiorstwie społecznym pomiar wpływu pełni dodatkową funkcję: jest mechanizmem ochrony misji przed dryfem ekonomicznym. Nie służy on wyłącznie mierzeniu sukcesu społecznego, lecz pozwala dostrzec, czy wzrost finansowy nie zmienia stopniowo populacji odbiorców, produktu i praktyk w sposób odsuwający organizację od jej pierwotnego celu.

Autor książki opisuje impact investing jako połączenie kapitału finansowego z oczekiwaniem mierzalnego rezultatu społecznego, sugerując zarządzanie wpływem na poziomie całego portfela. W tym sektorze niezbędna jest precyzyjna terminologia, gdyż „ESG investing", „sustainable investing" i „impact investing" nie są pojęciami tożsamymi. GIIN definiuje impact investments jako inwestycje dokonywane z intencją generowania pozytywnego, mierzalnego wpływu społecznego lub środowiskowego obok zwrotu finansowego. Do kluczowych elementów zalicza intentionality, wykorzystanie dowodów i danych przy projektowaniu inwestycji oraz zarządzanie performance wpływu. Tym samym impact nie jest tutaj wyłącznie charakterystyką aktywa ex post – powinien znajdować się w samej tezie inwestycyjnej. Pozwala to odróżnić inwestora kupującego udziały w

przedsiębiorstwie już wywierającym pozytywny wpływ, od takiego, którego kapitał przyczynia się do zwiększenia tego wpływu.

Pierwsza sytuacja mówi o właściwościach aktywa; druga dotyczy additionality inwestora. Impact Europe mocno podkreśla tę różnicę, rozdzielając impact na poziomie investee, wkład inwestora w rozwój investee oraz jego wkład w zmianę całego ekosystemu. Jest to istotne, gdyż fundusz może raportować wszystkie społeczne rezultaty spółki portfelowej jako własny sukces, mimo że jego kapitał nie zmienił istotnie trajektorii przedsiębiorstwa. W świetle wcześniejszej analizy byłby to klasyczny problem overclaimingu attribution.

Dlatego kluczowe pytanie w impact investingu brzmi nie tylko: „czy spółka robi dobro?", lecz przede wszystkim: „co zmieniło się dlatego, że właśnie ten kapitał został dostarczony w tej formie, w tym momencie i z takim wsparciem?". Inwestor może wnosić dodatkowość poprzez kapitał o wyższym ryzyku, dłuższy horyzont czasowy, wsparcie organizacyjne, governance, otwieranie sieci kontaktów, budowę systemu impact management lub katalizowanie innych źródeł finansowania. GIIN wskazuje również, że pomiar i zarządzanie wpływem powinny obejmować skutki zarówno pozytywne, jak i negatywne, a odbywać się na poziomie portfela, nie tylko pojedynczej inwestycji. Perspektywa ta łączy się z poprzednią częścią analizy: fundusz nie powinien maksymalizować pojedynczego outcome kosztem koncentracji ryzyka, odporności portfela lub negatywnych externalities niewidocznych w głównym wskaźniku. Szczególną trudność stanowi moment wyjścia (exit).

Jeżeli inwestor buduje spółkę społeczną, a następnie sprzedaje ją podmiotowi ograniczającemu misję, krótkookresowy impact może wyglądać znakomicie, podczas gdy trwałość systemowa okazuje się słaba. Ocena impact investingu powinna zatem uwzględniać nie tylko wejście kapitału, ale również governance w czasie inwestycji, trwałość misji i konsekwencje wyjścia.

Fundacje i grantodawcy: podmiot, który mierząc innych, projektuje cały ekosystem bodźców

Fundatorzy zajmują w architekturze wpływu pozycję szczególną. Często sami nie wykonują interwencji, lecz poprzez konstrukcję konkursu, warunki finansowania, długość grantu, wskaźniki, zasady kwalifikowalności i rytm raportowania określają zachowanie organizacji wykonawczych. Autor książki słusznie podkreśla, że grantodawcy w znacznym stopniu nadają ton kulturze całego sektora. Z perspektywy teorii instytucjonalnej fundator nie jest więc tylko płatnikiem – jest architektem reguł gry. Jeśli finansuje wyłącznie projekty dwunastomiesięczne, premiuje rezultaty możliwe do wykazania w tym krótkim czasie.

Jeżeli nie finansuje kosztów danych, a jednocześnie wymaga ewaluacji, przenosi ciężar pomiaru na grantobiorcę. Jeśli akceptuje tylko z góry określone działania, ogranicza zdolność adaptacji. Gdy

każda porażka grozi utratą kolejnego grantu, tworzy racjonalny bodziec do ukrywania błędów. Fundator jest zatem częściowo odpowiedzialny za jakość wiedzy generowanej przez sektor; może żądać prawdy i jednocześnie projektować warunki, w których ta prawda jest ekonomicznie kosztowna. Właśnie dlatego Impact Europe traktuje wkład inwestora i grantodawcy szerzej niż sam transfer pieniędzy. Może on obejmować wzmacnianie stabilności finansowej, odporności organizacji, praktyk impact management i zdolności organizacyjnych podmiotów wspieranych. Koncepcja ta podważa potoczny podział kosztów na „dobre, programowe" i „złe, administracyjne". System danych, kompetentna księgowość, rozwój kadry, cybersecurity czy rezerwa płynności nie są bezpośrednim impactem, ale mogą stanowić kapitał produkcyjny wpływu.

Należy jednak zachować równowagę. Finansowanie rdzenne nie zwalnia z accountability, a wzrost zaufania nie oznacza rezygnacji z dowodów. Trafniejszym modelem jest trust plus verification: fundator nie mikrozarządza środkami, lecz wspólnie z organizacją definiuje outcomes, ryzyka, zasady uczenia się i zakres odpowiedzialności. W tym sektorze szczególnie ostro ujawnia się problem władzy epistemicznej. To często fundator decyduje, które outcomes są „ważne". Beneficjent może uważać za kluczową autonomię, organizacja – stabilność relacji, a fundator – zatrudnienie. Jeśli miernik fundatora staje się dominującym kryterium, pieniądz zaczyna kształtować definicję dobra. Dlatego stakeholder participation nie jest jedynie dodatkiem konsultacyjnym, lecz mechanizmem równoważenia asymetrii władzy pomiędzy płatnikiem i człowiekiem, którego życie jest przedmiotem interwencji.

Pomiar wpływu jako narzędzie odpowiedzialności i władzy

Autor książki proponuje efektowne rozróżnienie między „Sport Plus" a „Plus Sport". W pierwszym przypadku sport jest instrumentem prowadzącym do rezultatu społecznego – takiego jak zatrudnienie, edukacja, reintegracja czy bezpieczeństwo. W drugim zaś celem jest sam rozwój uczestnictwa sportowego. Heurystyka ta okazuje się przydatna przede wszystkim dlatego, że zapobiega błędowi polegającemu na finansowaniu „sportu jako narzędzia przeciwdziałania bezrobociu", by następnie raportować jedynie liczbę przeprowadzonych treningów. Należy jednak skorygować zbyt ostre przeciwstawienie tych dwóch podejść.

„Plus Sport", czyli uczestnictwo sportowe jako cel sam w sobie, nie jest społecznie bezwartościowe. Aktywność fizyczna może oddziaływać na zdrowie fizyczne i psychiczne, dobrostan, kapitał społeczny oraz rozwój jednostki nawet wtedy, gdy nie jest opakowana w odrębny program reintegracyjny. Sport England ocenia obecnie wartość społeczno-ekonomiczną sportu i aktywności

fizycznej poprzez rezultaty dotyczące zdrowia, wellbeing, rozwoju indywidualnego, społeczności i gospodarki. W odniesieniu do roku 2023/24 instytucja ta oszacowała łączną społeczną wartość community sport and physical activity w Anglii na 122,9 mld GBP. Nie oznacza to oczywiście, że każdemu treningowi można przypisać proporcjonalną część tej kwoty – szacunek makroekonomiczny nie stanowi attribution dla pojedynczego klubu. Pokazuje on jednak, że sam udział w sporcie może być nośnikiem wartości społecznej. Prawidłowe rozróżnienie powinno więc brzmieć inaczej.

Jeśli celem programu jest sport, to frekwencja, regularność aktywności, zdrowie i doświadczenie uczestnika mogą być adekwatnymi outcomes. Jeśli natomiast sport został użyty jako środek do ograniczania bezrobocia lub przestępczości, sama liczba treningów nie wystarcza jako dowód realizacji celu. Należy mierzyć rezultat, którego finansowanie miało dotyczyć. To subtelna, lecz istotna korekta: nie istnieje hierarchia, w której „Sport Plus" jest z definicji dojrzały, a „Plus Sport" prymitywny. Istnieje jedynie obowiązek zgodności wskaźnika z teorią celu.

Sport stanowi również dobry przykład problemu mediacji. Aktywność może wpływać na zatrudnienie nie bezpośrednio, lecz poprzez zwiększenie poczucia własnej skuteczności, samodyscypliny, relacji społecznych, zdrowia lub zaufania. Program mierzący wyłącznie ostateczne zatrudnienie może zatem nie rozumieć mechanizmu zmiany, mimo słusznego skoncentrowania się na rezultacie końcowym. Właśnie tutaj przydatna jest pełna Theory of Change: outcome końcowy i mediatory nie są konkurencyjnymi miernikami, lecz różnymi poziomami tego samego modelu.

Porównanie sektorów pozwala sformułować ważniejszą tezę: impact measurement jest zawsze częścią systemu accountability. Nie istnieje on poza relacjami władzy. To, kto żąda danych, wpływa na to, jakie dane powstają. W korporacji trafiają one do zarządu, inwestorów, regulatora, pracowników i rynku. W administracji – do organów politycznych, kontroli, parlamentu, obywateli i audytorów. W NGO – do rady, fundatora, społeczności i beneficjentów. W impact investing – do inwestorów kapitałowych, komitetu inwestycyjnego, spółek portfelowych i końcowych odbiorców wpływu. W fundacji – do rady powierniczej, donatorów i organizacji finansowanych. W sporcie – do uczestników, sponsorów, samorządów, federacji i społeczności. Nie są to neutralni odbiorcy informacji; każdy z nich dysponuje inną władzą sankcyjną.

Inwestor może wycofać kapitał. Fundator może nie przedłużyć grantu. Obywatel może głosować, choć jego indywidualny wpływ jest rozproszony. Beneficjent może zrezygnować z programu, lecz często nie posiada alternatywy. Dlatego architektura impact measurement jest zarazem architekturą tego, czyje pytania otrzymują pierwszeństwo.

Adekwatność metody zamiast uniwersalnego standardu pomiaru

Socjologia organizacji wprowadza tutaj użyteczne pojęcie izomorfizmu instytucjonalnego. Organizacje mają bowiem skłonność do upodabniania się do podmiotów uznawanych za profesjonalne, legitymizowane lub kontrolujące zasoby. W efekcie sektor NGO przejmuje język biznesowych KPI, fundatorzy – terminologię venture capital, administracja – zasady zarządzania przedsiębiorstwem, a firmy – retorykę social impact.

Takie zapożyczenia bywają produktywne, lecz mogą również prowadzić do category errors. Organizacja społeczna nie jest startupem z gorszym modelem biznesowym. Administracja nie jest przedsiębiorstwem należącym do podatników w takim samym sensie, w jakim spółka należy do akcjonariuszy. Beneficjent nie zawsze jest klientem, gdyż często nie posiada ani wyboru, ani siły nabywczej. Fundacja nie staje się VC tylko dlatego, że tworzy „portfolio grantów". Autor książki przywołuje tu uwagi OECD dotyczące niedopasowania istniejących metod do social economy: wiele systemów stworzono pod potrzeby fundatorów i sektora for-profit, podczas gdy organizacje społeczne mierzą zjawiska takie jak inkluzja, zaufanie, community development czy wellbeing – obszary szczególnie trudne do prostego skwantyfikowania. Z tego wynika kluczowa zasada: transferuj pytania, niekoniecznie instrumenty. Biznes może nauczyć NGO dyscypliny w alokacji kapitału, ale nie musi narzucać mu kwartalnego rytmu wyniku. Sektor społeczny może z kolei nauczyć biznes partycypacji interesariuszy, administracja wnieść legalność i procedural fairness, venture capital – logikę portfela i eksperymentowania, a nauka – rygor wnioskowania przyczynowego.

Żaden sektor nie posiada modelu kompletnego. Wśród prób unifikacji najbardziej obiecująca wydaje się nie standaryzacja wszystkich wskaźników, lecz zasada proporcjonalności. OECD definiowała ją jako równoważenie kosztów i korzyści pomiaru w przedsiębiorstwach społecznych. W nowszych materiałach OECD oraz Komisja Europejska podtrzymują podejście, według którego system pomiarowy powinien być dostosowany do misji, skali, potrzeb decyzyjnych i możliwości organizacji. Proporcjonalność oznacza, że fundacja rozdysponowująca niewielkie granty nie powinna obarczać każdego projektu ciężarem quasi-eksperymentalnej ewaluacji. Z drugiej strony program publiczny obejmujący miliony obywateli i miliardy złotych powinien podlegać znacznie silniejszemu rygorowi niż lokalny pilotaż. Fundusz impact investing, dokonujący dużej alokacji, może racjonalnie inwestować w zaawansowany system IMM, podczas gdy małe stowarzyszenie może polegać na kilku dobrze dobranych outcomes i jakościowej interpretacji. Rygor nie jest więc

ilością metod, lecz proporcją między znaczeniem decyzji, kosztem błędu i jakością wymaganego dowodu.

Również sposób myślenia przyczynowego powinien być dostosowany do specyfiki sektora. W programie klinicznym możliwy jest RCT; w krajowej reformie podatkowej – nie. W kampanii rzeczniczej bardziej adekwatny może być process tracing, natomiast w przedsiębiorstwie wpływ produktu można badać poprzez quasi-eksperyment, dane panelowe albo natural experiment. W impact investing konieczne jest oddzielenie impactu spółki od contribution inwestora, a w philanthropy należy dodatkowo pytać, czy grant był dodatkowy wobec innych źródeł finansowania. Nie istnieje więc „złoty standard" identyczny dla wszystkich sektorów. Istnieje raczej złoty standard adekwatności pytania i metody.

W przedsiębiorstwie maksymalizacja oznacza znalezienie konfiguracji, w której impact zostaje zintegrowany z tworzeniem wartości ekonomicznej i ograniczeniami finansowymi. W państwie musi być ona podporządkowana prawu, demokratycznej legitymizacji i równości. W NGO oznacza przede wszystkim maksymalizowanie realizacji misji przy jednoczesnej ochronie przed dyktatem fundatora. W przedsiębiorstwie społecznym wymaga równowagi pomiędzy zdolnością ekonomiczną a mission lock. W impact investing obejmuje zarządzanie całym portfelem wpływu, financial return i additionality inwestora. W fundacji jest w dużej mierze problemem projektowania bodźców u innych organizacji. W sporcie zaś zależy od tego, czy sport jest dobrem finalnym, czy technologią społecznej zmiany.

Uniwersalność procesu a specyfika języka instytucjonalnego

Oznacza to, że triada Stevenson-Dodda (Understand-Measure-Maximise) może być uniwersalna jedynie na wysokim poziomie abstrakcji; jej operacjonalizacja musi mieć charakter sektorowy. Można wręcz stwierdzić, że dojrzałość sektora najłatwiej rozpoznać po tym, komu i na jakiej podstawie instytucja ma prawo odmówić. Firma może zrezygnować z nieopłacalnego segmentu rynku, o ile przestrzega prawa antydyskryminacyjnego. NGO może ograniczyć grupę beneficjentów zgodnie ze swoim celem (purpose), fundusz może odmówić inwestycji, a fundacja – przyznania grantu. Państwo dysponuje znacznie mniejszą swobodą, gdyż obowiązek świadczenia usług lub równego traktowania wynika z przepisów prawa. To właśnie tutaj ekonomika maksymalizacji zderza się z publicznym obowiązkiem. Organ publiczny nie może po prostu skierować wszystkich środków do tych obywateli, u których koszt osiągnięcia rezultatu (outcome) jest najniższy. Jeśli świadczenie ma charakter prawa, niektóre formy „efektywności" są normatywnie niedopuszczalne.

Różnice sektorowe nie są zatem szczegółem organizacyjnym, lecz wyznaczają granicę stosowalności samego języka optymalizacji.

Pomimo tych rozbieżności można wskazać wspólną strukturę dojrzałego zarządzania wpływem. Nie wymaga ona identycznych wskaźników KPI, lecz jasnego celu, przejrzystej Teorii Zmiany (Theory of Change), proporcjonalnego pomiaru oraz uczciwego rozróżnienia między wkładem (contribution) a przypisaniem sukcesu (attribution). Wymaga także wykorzystywania danych w procesach decyzyjnych i gotowości do ujawniania negatywnych bądź niezamierzonych skutków. Co znamienne, wspólny przewodnik OECD i Komisji Europejskiej dla ekonomii społecznej z 2024 roku również posługuje się logiką Measure, Manage and Maximise, podkreślając translację misji na zmianę, holistyczne ujmowanie konsekwencji, proporcjonalność, orientację na interesariuszy, transparentność i ciągłe doskonalenie.

Nie dowodzi to zależności od koncepcji Stevenson-Dodda, lecz pokazuje, że jego pragmatyczne intuicje współbrzmią z ewolucją międzynarodowego podejścia do impact management. Kluczowe jest jednak to, czego wspólny model nie powinien robić: Nie powinien kolonizować różnorodnych instytucji jednym językiem efektywności. Administracja nie może utracić legalności, NGO głosu beneficjenta, przedsiębiorstwo odpowiedzialności za własny core business, inwestor kwestii dodatkowości (additionality), fundator świadomości własnej władzy, a sport wartości uczestnictwa samego w sobie. Prawdziwa dojrzałość wpływu nie polega więc na ujednoliceniu wszystkich sektorów, lecz na tym, że każdy z nich potrafi odpowiedzieć we własnym instytucjonalnym języku na kluczowe pytania: jaki stan świata ma się zmienić dzięki naszemu istnieniu, jakie mamy dowody na tę zmianę, jaki jest nasz rzeczywisty wkład i co powinniśmy zrobić inaczej, gdy odpowiedź okazuje się mniej korzystna, niż oczekiwaliśmy. Na tym poziomie można przejść od porównania sektorów do zasadniczego problemu polskiego systemu.

Instytucjonalna obecność mierników nie gwarantuje zarządzania wpływem

Warto zbadać, czy i w jakim sensie polska architektura instytucjonalna sprzyja kulturze „Do". Należy przeanalizować, jak na sposób mierzenia wpływu oddziałują legalizm administracyjny, budżet zadaniowy, dotacje projektowe, fundusze europejskie, system kontroli, zamówienia publiczne, kultura audytu, społeczeństwo obywatelskie oraz relacja fundator-grantobiorca. Dopiero taka analiza pozwoli odpowiedzieć w sposób empiryczny, a nie hasłowy, jak jest w Polsce, dlaczego tak się dzieje i gdzie rzeczywiście istnieje przestrzeń przejścia od sprawozdawania działania do zarządzania różnicą.

Zastosowanie aparatu Stevensona-Dodda do warunków polskich wymaga powściągliwości w stawianiu prostej diagnozy, jakoby Polska była po prostu systemem typu „Do". Takie twierdzenie byłoby zbyt szerokie empirycznie i zbyt ubogie teoretycznie.

Polski sektor publiczny, system funduszy europejsych, organizacje społeczne i instytucje ekonomii społecznej dysponują rozwiniętymi mechanizmami monitorowania, miernikami rezultatów, wyspecjalizowanymi strukturami ewaluacyjnymi, procedurami wdrażania rekomendacji oraz rosnącymi zasobami danych. Jednocześnie istnieją mocne przesłanki, że samo wytwarzanie informacji o rezultatach nie zawsze oznacza, iż stają się one podstawowym kryterium realokacji zasobów i korekty działania. Właściwym przedmiotem polskiej diagnozy nie jest zatem alternatywa „mierzymy – nie mierzymy", lecz relacja między czterema procesami: wytwarzaniem danych, ich interpretacją, umiejscowieniem w strukturze decyzyjnej oraz rzeczywistą zdolnością do zmiany programu. Rozróżnienie to ma znaczenie zasadnicze, gdyż Polska posiada już wiele instytucjonalnych elementów charakterystycznych dla modelu „Difference".

Ustawa o zasadach prowadzenia polityki rozwoju umocowuje ewaluację jako element tej polityki; zadania koordynacyjne obejmują tu prowadzenie ewaluacji strategii i instrumentów, opracowywanie standardów oraz formułowanie rekomendacji dla programowania i realizacji działań. W polityce spójności funkcjonuje Krajowa Jednostka Ewaluacji, która koordynuje proces ewaluacyjny, monitoruje plany, jakość badań i wdrażanie rekomendacji, a także zarządza Bazą Badań Ewaluacyjnych oraz Systemem Wdrażania Rekomendacji. Fakt ten uniemożliwia postawienie tezy, że polskie państwo nie posiada instytucjonalnej kultury ewaluacyjnej. Posiada ją, lecz jest ona nierównomiernie rozwinięta sektorowo i nierównomiernie powiązana z centrami alokacji zasobów.

Dobrym laboratorium tego napięcia jest budżet zadaniowy. W 2026 r. Ministerstwo Finansów nadal publikuje katalog funkcji, zadań, podzadań i działań budżetowych wraz ze zbiorczym zestawieniem celów i mierników. Oznacza to, że państwowa architektura finansowa formalnie dysponuje językiem wykraczającym poza klasyfikację ekonomiczną wydatków: próbuje wiązać zasób z celem i miernikiem. Istnienie takiej struktury jest istotne, gdyż tworzy warunek konieczny dla zarządzania przez rezultaty, choć nie jest on warunkiem wystarczającym.

Najwyższa Izba Kontroli w kontroli poświęconej planowaniu i realizacji zadaniowego układu wydatków sformułowała wcześniej ocenę zdecydowanie krytyczną: budżet zadaniowy nie odgrywał na poziomie centralnym znaczącej roli w zarządzaniu środkami publicznymi i nie doprowadził do oczekiwanej poprawy efektywności, rozliczalności ani komunikacji z interesariuszami.

Decoupling i dominacja legalizmu nad optymalizacją wpływu

Chronologia ma tutaj znaczenie. Ustalenia te nie mogą zostać mechanicznie przeniesione na rok 2026 i przedstawione tak, jakby późniejsze reformy niczego nie zmieniły. Pokazują one jednak zjawisko kluczowe dla przyjętej teorii: instytucja może posiadać cele i mierniki, nie posiadając jeszcze budżetu rzeczywiście sterowanego przez cele i mierniki. W języku teorii instytucjonalnej mówimy tu o decouplingu, czyli częściowym oddzieleniu struktury formalnej od praktyki decyzyjnej. Meyer i Rowan wskazywali, że organizacje przyjmują struktury uznawane za racjonalne i legitymizujące, lecz nie każda z nich w równym stopniu przenika do codziennego działania.

Polski budżet zadaniowy warto badać właśnie z tej perspektywy: nie pytając jedynie o istnienie mierników, lecz sprawdzając, czy ich wynik realnie wpływa na przyszłą alokację środków, kształt programu lub decyzję o dezinwestycji. Jeśli działanie w kolejnym roku otrzymuje podobne finansowanie niezależnie od stopnia realizacji outcome, miernik pełni głównie funkcję sprawozdawczą. Dopiero gdy niski wynik uruchamia analizę mechanizmu, korektę Theory of Change, zmianę instrumentu lub przesunięcie środków, zaczyna on pełnić funkcję zarządczą. Ta różnica odpowiada Stevenson-Doddowskiemu przejściu od retrospektywnej ewaluacji do impact management.

Nie można przy tym sprowadzać problemu do rzekomej niechęci polskiej administracji do efektywności. Coroczna kontrola wykonania budżetu prowadzona przez NIK obejmuje bowiem nie tylko zgodność finansową, ale i kontrolę wykonania zadań. W odniesieniu do 2025 roku NIK przeprowadziła ponad sto takich kontroli w administracji rządowej i innych jednostkach korzystających ze środków publicznych. Polska infrastruktura accountability obejmuje zatem pytania o wykonanie. Problem jest jednak subtelniejszy: kontrola wykonania nie stanowi jeszcze systemu ciągłej optymalizacji wpływu.

Wyjaśnienie polskiej specyfiki wymaga uwzględnienia legalistycznej tradycji administracji. Państwo prawa musi kłaść duży nacisk na legalność wydatku, kompetencje organu, prawidłowość postępowania, zachowanie procedur i udokumentowanie podstawy decyzji. Nie jest to biurokratyczna aberracja – formalizacja chroni obywatela przed arbitralnością władzy, ogranicza klientelizm, ułatwia kontrolę sądową oraz audyt i zabezpiecza równe stosowanie prawa. Z perspektywy impact management pojawia się tu jednak asymetria poznawcza. Legalność jest względnie łatwo audytowalna, wpływ społeczny - znacznie trudniej. Można bowiem jednoznacznie ustalić, czy wydatek był zgodny z planem finansowym, czy przeprowadzono wymagane postępowanie i czy dokumentacja jest kompletna.

Znacznie trudniej jednak rozstrzygnąć, czy program zmienił zachowania ludzi, w jakim stopniu zmiana ta była dodatkowa, czy nie wystąpiło zjawisko displacement i czy rezultat utrzymał się przez pięć lat. System kontroli ma zatem naturalną skłonność do preferowania tych aspektów rzeczywistości, które pozostawiają jednoznaczny ślad dokumentacyjny. Nie wynika to z braku zainteresowania skutecznością ze strony kontrolerów, lecz z faktu, że technologia kontroli kształtuje ontologię tego, co łatwo uznać za fakt administracyjny.

Od rytuału dokumentacji do audytowalności epistemicznej

Można tu przywołać Mertonowskie pojęcie **goal displacement**. Procedura, stworzona jako środek do realizacji celu, może z czasem zyskać autonomiczną rangę. Organizacja zaczyna być oceniana przez pryzmat prawidłowego wykonania procedury, więc racjonalnie inwestuje energię właśnie w ten proces. W skrajnym przypadku społeczny cel interwencji pozostaje jedynie deklarowanym sensem działania, podczas gdy operacyjnym celem staje się przejście audytu.

To jednak punkt, w którym krytyka Stevensona-Dodda wymaga aksjologicznej korekty. Nie wolno przeciwstawiać zasady **Function First** legalności tak, jakby rezultat społeczny legitymizował dowolną procedurę. Państwo może osiągnąć outcome i jednocześnie naruszyć prawa jednostki. Administracja typu "Difference" musi być zatem efektywna **w granicach normatywnych**, a nie zamiast nich. Michael Power opisywał rozwój współczesnych instytucji poprzez pojęcie **audit society**: coraz większe obszary działania poddawane są formalnej weryfikacji, lecz zdolność audytu zależy od wcześniejszego skonstruowania rzeczywistości w formie audytowalnej.

To spostrzeżenie wyjątkowo trafnie pomaga interpretować polski system dotacyjny i publiczny. Jeśli rezultat ma być kontrolowalny, organizacja potrzebuje dokumentu. Umowa, lista obecności, faktura, protokół, zaświadczenie i karta czasu pracy dają silniejsze bezpieczeństwo dowodowe niż opowieść o odzyskanej sprawczości człowieka. W konsekwencji to, co łatwo udokumentować, zyskuje przewagę nad tym, co społecznie ważne. Stevenson-Doddowska zasada **auditable outcomes** ma więc w polskim kontekście charakter podwójny. Z jednej strony jest radykalnie potrzebna, gdyż pozwala wyjść poza deklaratywne success stories. Z drugiej – jeśli pojęcie audytowalności zostanie zawężone do dokumentu administracyjnego, może ponownie skierować system ku outputs. Trwała poprawa dobrostanu nie musi pozostawiać faktury. Wzrost zaufania społecznego nie posiada świadectwa odbioru. Godność nie generuje protokołu wykonania.

Potrzebna jest zatem szersza koncepcja **audytowalności epistemicznej**: outcome powinien posiadać dowód możliwy do niezależnej kontroli, ale takim dowodem może być również zwalidowana skala, rejestr administracyjny, rzetelnie udokumentowany wywiad, analiza kontrfaktyczna, badanie panelowe albo triangulacja kilku źródeł. Najmocniejszym kontrargumentem wobec tezy o dominacji kultury "Do" jest polski system ewaluacji polityki spójności.

Wytyczne na lata 2021-2027 ustanawiają rozbudowaną architekturę instytucjonalną, obejmującą Krajową Jednostkę Ewaluacji, jednostki ewaluacyjne, plany ewaluacji, monitoring jakości badań, system rekomendacji i bazę badań. Jest to infrastruktura o stopniu formalizacji znacznie przewyższającym praktykę wielu pojedynczych organizacji społecznych. Co szczególnie istotne, polskie ewaluacje nie ograniczają się do prostego monitoringu produktów. Dostępne badania polityki spójności wykorzystują **theory-based evaluation, grupy kontrolne, analizy kontrfaktyczne, metody statystyczne, wywiady, ankiety, studia przypadku i triangulację**. W ewaluacji wpływu polityki spójności 2014-2020 na system kształcenia i szkolenia zastosowano równolegle ewaluację opartą na teorii oraz analizę wykorzystującą grupy kontrolne. W badaniu wsparcia działalności B+R przedsiębiorstw użyto analiz kontrfaktycznych, pozwalających oszacować, jak aktywność badawczo-rozwojowa beneficjentów wyglądałaby bez interwencji. Nie mamy więc do czynienia z państwem, które metodologicznie nie zna attribution, kontrfaktyczności czy Theory of Change. Przeciwnie – istnieją w nim zaawansowane wyspy kompetencji ewaluacyjnej.

Właściwe pytanie brzmi jednak: **jak szeroko wiedza tego typu przenika poza ekosystem funduszy europejskich oraz jak silnie oddziałuje na kolejne alokacje polityczne?** Sam System Wdrażania Rekomendacji jest próbą rozwiązania właśnie tego problemu, ponieważ pozwala monitorować stan wdrożenia zaleceń wynikających z badań.

Spending completion zamiast mission completion

Podejście to jest bardzo bliskie koncepcji zamkniętej pętli uczenia przedstawionej wcześniej. Polski przypadek sugeruje zatem nie deficyt ewaluacji, lecz możliwy **problem translacji ewaluacji na governance**. Państwo potrafi wytwarzać wiedzę o wpływie; znacznie trudniej jest jednak sprawić, by wiedza ta konkurowała z presją polityczną, historycznymi budżetami, harmonogramami absorpcji, zobowiązaniami instytucjonalnymi czy cyklem wyborczym. Fundusze europejskie są w tym kontekście szczególnie interesujące, ponieważ jednocześnie rozwijają kulturę impact evaluation i utrwalają silną logikę wykonania finansowego. Aktualny Portal Funduszy Europejskich publikuje szczegółowe, stale aktualizowane dane o liczbie umów, wnioskach o płatność, formach

prawnych beneficjentów i miejscach realizacji interwencji. Choć niesie to ogromną wartość informacyjną, obnaża też istotne rozróżnienie: **stan wdrażania nie jest jeszcze stanem wpływu**. Wydatkowanie środków odpowiada jedynie na pytanie, czy instrument został uruchomiony.

Liczba umów mówi o skali, liczba beneficjentów o zasięgu, a wskaźnik produktu o tym, co wytworzono. Dopiero rezultat i impact pozwalają stwierdzić, czy rzeczywistość uległa zmianie. Przykłady z badań ewaluacyjnych polityki spójności pokazują, że polski system dostrzega tę różnicę. W jednej z syntez ewaluacji dotyczących COVID-19 wysokie wykorzystanie alokacji i osiąganie wskaźników produktu analizowano oddzielnie od pytania o skuteczność interwencji w poszczególnych sektorach oraz jej szersze konsekwencje. To rozróżnienie powinno stać się zasadą całego zarządzania publicznego: **100% wykonania budżetu nie oznacza 100% wykonania celu**. W polskiej debacie administracyjnej „niewykonanie wydatków" często kojarzy się jednoznacznie negatywnie, jednak z perspektywy impact management sytuacja jest bardziej złożona.

Jeśli program osiągnął cel przy mniejszych nakładach, niewydanie całej kwoty może świadczyć o efektywności. Jeśli zaś środki wydano w całości, a rezultat nie nastąpił, pełne wykonanie finansowe nie jest sukcesem społecznym. System nagradzający wykonanie alokacji niezależnie od końcowego wpływu tworzy bodziec do **spending completion zamiast mission completion**. Drugim kluczowym elementem polskiej architektury jest **projektification**, czyli organizowanie coraz większej części aktywności społecznej poprzez projekty o określonym początku, końcu, zamkniętym budżecie, miernikach i harmonogramie. Model projektowy ma ogromne zalety: umożliwia konkurencję, porównanie ofert, kontrolę kosztów, testowanie nowych działań i przypisanie odpowiedzialności. Jego wada ujawnia się jednak wtedy, gdy problem społeczny ma charakter ciągły, a finansowanie – epizodyczny. Bezdomność nie kończy się 31 grudnia tylko dlatego, że wygasł okres kwalifikowalności wydatków. **Trauma nie ma harmonogramu projektowego**.

Kapitał społeczny buduje się latami. Relacja z młodzieżą zagrożoną wykluczeniem może wymagać długiego procesu budowania zaufania, którego nie da się restartować wraz z każdym nowym konkursem. Projektowość prowadzi więc do **temporal mismatch** między cyklem finansowania a cyklem zmiany społecznej. Program zaczyna się w momencie podpisania umowy i kończy wraz z grantem, podczas gdy outcome może wystąpić znacznie później. Jeśli system nie finansuje obserwacji po zakończeniu projektu, organizacja uczy się mierzyć tylko to, co mieści się w ramach czasowych finansowania. W ten sposób czas staje się ukrytą determinantą miary: krótkie finansowanie preferuje krótkie outcomes.

Zależność zasobowa a horyzont poznawczy organizacji

Sektor organizacji pozarządowych stanowi istotny komponent infrastruktury społecznej państwa, a nie marginalną grupę hobbystycznych stowarzyszeń. Jednocześnie badanie „Kondycja organizacji pozarządowych 2024", przeprowadzone przez Klon/Jawor na reprezentatywnej próbie 1012 aktywnych podmiotów, obnaża skalę problemów zasobowych. Aż 72% badanych organizacji deklarowało trudności w zdobywaniu funduszy, 51% problemy z płynnością finansową, 55% nadmierną biurokrację i formalności, a 49% niejasność przepisów regulujących ich działalność.

Te dane są kluczowe w perspektywie Stevenson-Dodda, gdyż pokazują, że żądanie od organizacji „lepszego impact measurement" nie może abstrahować od ich capacity to measure. Podmiot walczący o płynność, oparty na okresowych projektach i niewielkim zespole, często nie dysponuje analitykiem, systemem CRM, możliwością prowadzenia follow-up ani środkami na kontrfaktyczną ewaluację. Według tych samych badań w 2023 r. rezerwę finansową posiadało jedynie 35% stowarzyszeń i fundacji; środki publiczne (krajowe i zagraniczne) stanowiły około 44% budżetu badanego sektora.

Jest to strukturalnie istotne, ponieważ płynność wpływa nie tylko na bezpieczeństwo organizacji, ale bezpośrednio na horyzont poznawczy jej zarządu. Instytucja zagrożona utratą finansowania w ciągu trzech miesięcy racjonalnie poświęca więcej uwagi kolejnemu konkursowi niż pięcioletniemu badaniu trwałości outcome. Mamy tu do czynienia z formą organizacyjnej krótkowzroczności, która nie musi wynikać ze słabego przywództwa, lecz może być racjonalną adaptacją do środowiska finansowania.

„Grantoza" jako instytucjonalny mechanizm kształtowania zachowań Polska debata pozarządowa posługuje się nieformalnym pojęciem „grantozy", aby opisać zależność organizacji od kolejnych projektów dotacyjnych. Nauki społeczne pozwalają przełożyć tę intuicję na bardziej analityczny język resource dependence. Podmiot zależny od zewnętrznego zasobu dostosowuje swoje zachowanie do wymogów dysponenta tego zasobu. Jeśli finansowanie otrzymuje się na konkretny projekt, racjonalne staje się tworzenie projektów. Jeśli konkurs definiuje grupę docelową, organizacja zaczyna posługiwać się właśnie tą kategorią. Jeśli wskaźnik konkursowy mierzy liczbę wydarzeń, to wydarzenia zyskują znaczenie operacyjne. Jeśli koszty administracyjne są traktowane podejrzliwie, organizacja zaniża overhead lub przenosi go do innych pozycji. W ten sposób instrument finansowania współtworzy ontology of action: określa, jakie jednostki aktywności zostają uznane za realne.

Badania Klon/Jawor wskazują jednocześnie na pewną dywersyfikację: przeciętna organizacja korzystała w 2023 r. z kilku rodzajów przychodów, a udział środków publicznych w budżetach

sektora spadł względem wcześniejszych pomiarów. Nie należy więc opisywać całego sektora jako całkowicie zależnego od jednego modelu grantowego. Problemem pozostaje raczej fakt, że projektowość jest szczególnie wpływowa w tych organizacjach, które realizują zadania publiczne i zatrudniają stały personel – czyli tam, gdzie stabilność instytucjonalna jest najbardziej potrzebna.

Ważnym kontrprzykładem jest Rządowy Program Rozwoju Organizacji Obywatelskich. W edycji 2025 Priorytet 1a PROO przewiduje możliwość finansowania działań misyjnych i wzmacniania instytucjonalnego organizacji. Samo istnienie takiego instrumentu jest koncepcyjnie istotne: stanowi przyznanie, że skuteczność organizacji nie zależy wyłącznie od bezpośrednio finansowanych outputów, lecz również od jej zdolności, infrastruktury i stabilności.

Podobnie NIW prowadzi ewaluacje swoich programów. W 2024 r. zrealizowano między innymi ewaluację mid-term PROO oraz bieżące ewaluacje programów NOWEFIO i innych instrumentów. W kwietniu 2025 r. Komitet Sterująco-Monitorujący NOWEFIO omawiał bieżącą ewaluację programu, a w sierpniu – sprawozdanie za 2024 r. i możliwe zmiany regulaminu kolejnej edycji. Jest to właśnie instytucjonalny zalążek feedback loop: badanie nie kończy się w archiwum, lecz trafia na forum dyskutujące o przyszłej konstrukcji instrumentu. Nie przesądza to oczywiście, w jakiej mierze rekomendacje faktycznie zmieniają kolejne decyzje – do wykazania pełnej pętli potrzebne byłoby prześledzenie konkretnych zaleceń i późniejszych zmian regulaminowych. Samo istnienie tego mechanizmu pokazuje jednak, dlaczego należy odrzucić uproszczoną narrację o Polsce jako jednolitym systemie proceduralnym.

Podobny kontrprzykład stanowi Krajowy Program Rozwoju Ekonomii Społecznej do 2030 r. Program ma charakter operacyjno-wdrożeniowy i wiąże interwencję w ekonomii społecznej z polityką społeczną oraz rozwojem lokalnym. Jego system przewiduje wskaźniki rezultatów, okresowy monitoring i możliwość reagowania na zmianę warunków. Dokument programowy wskazuje wręcz na coroczne przeglądy realizacji, ewaluację i możliwość redefiniowania wybranych elementów strategii wobec nowych wyzwań. Dodatkowo istnieje Zintegrowany System Monitorowania Sektora Ekonomii Społecznej, korzystający z danych statystyki publicznej i źródeł administracyjnych, co umożliwia porównywanie jednostek samorządu oraz śledzenie wskaźników programowych. Konstrukcja ta jest bliska Live Data nie w sensie sekundowej aktualizacji, lecz jako próba stworzenia trwałej infrastruktury informacyjnej zamiast jednorazowego raportu.

Psychologia i bariery systemowe w zarządzaniu wpływem

Część wskaźników ekonomii społecznej ma charakter skalowy: liczba podmiotów, przedsiębiorstw społecznych czy osób zatrudnionych. Są one niezbędne, lecz nie odpowiadają na wszystkie pytania o społeczną additionality tych podmiotów. Sam wzrost liczby przedsiębiorstw społecznych nie dowodzi bowiem, że poprawiła się sytuacja osób zagrożonych wykluczeniem lub jakość usług społecznych. Dojrzały system powinien zatem łączyć capacity indicators z beneficiary outcomes.

Również systemu zamówień publicznych nie można przedstawiać jako czysto proceduralnego mechanizmu ceny i formalności. Art. 242 Prawa zamówień publicznych pozwala stosować kryteria jakościowe odnoszące się m.in. do dostępności, potrzeb użytkowników, aspektów społecznych, integracji zawodowej i społecznej, środowiska oraz innowacyjności. Z kolei art. 96 umożliwia ustanawianie wymagań związanych z realizacją zamówienia w tych samych obszarach. Prawo tworzy zatem instrumentarium pozwalające zastosować logikę Function First, choć nie gwarantuje to jakości ich wykorzystania. Klauzula zobowiązująca wykonawcę do zatrudnienia określonej grupy jest istotnym narzędziem społecznym, ale nadal może być outputem regulacyjnym, jeśli nie badamy później trwałości zatrudnienia, sytuacji pracownika lub dodatkowości tego wymagania. Kryterium społeczne bywa sprowadzane do kilku punktów w postępowaniu, co nie wpływa realnie na końcowy wybór. To doskonały przykład różnicy pomiędzy social procurement a social impact procurement.

Pierwsze z nich włącza do zamówienia wymaganie społeczne; drugie stara się ustalić, czy to wymaganie rzeczywiście wywołało zamierzoną zmianę. Nie oznacza to, że każda umowa publiczna wymaga rozbudowanej ewaluacji – zgodnie z zasadą proporcjonalności wymogi dowodowe powinny zależeć od wartości, ryzyka i społecznego znaczenia zamówienia. Jednak w największych kontraktach przesunięcie pytania z „czy wykonawca spełnił klauzulę?” na „jaką zmianę społeczną wytworzyła realizacja klauzuli?” stanowi naturalne rozwinięcie modelu Stevenson-Dodda.

Szczególnie ważną perspektywę wnosi tu teoria street-level bureaucracy Michaela Lipsky'ego. Polityka publiczna nie jest wykonywana mechanicznie, zgodnie z literą ustawy. Nauczyciele, pracownicy socjalni, urzędnicy, policjanci czy doradcy zawodowi dysponują pewnym zakresem dyskrecjonalności. Ograniczenia czasowe i zasobowe zmuszają ich do tworzenia praktycznych heurystyk, które współkształtują realny rezultat polityki. Z perspektywy impact management oznacza to, że centralnie zaprojektowany KPI może wywoływać odmienne zachowania na poziomie pierwszej linii. Urząd może formalnie maksymalizować terminowość, podczas gdy pracownicy wypracują praktykę ograniczającą liczbę trudnych spraw. System aktywizacyjny premiujący

skuteczne zatrudnienia może z kolei skłonić doradców do koncentracji uwagi na uczestnikach mających największe szanse na sukces.

Dlatego polska transformacja od „Do" do „Difference" nie może opierać się wyłącznie na zmianie wskaźników centralnych. Musi uwzględnić mikroekonomię i psychologię reakcji ludzi na wskaźniki. Jest to kolejny argument przeciwko przekonaniu, że wystarczy rozliczać z efektów. Outcome-based accountability jest dojrzała tylko wtedy, gdy projektant potrafi przewidzieć zachowania adaptacyjne powstające pod wpływem kontraktu. W administracji i polityce działa dodatkowo mechanizm blame avoidance: korzyść z innowacyjnego programu jest często rozproszona i odroczona, natomiast spektakularna porażka generuje natychmiastowy koszt polityczny. Ta asymetria skłania racjonalnego decydenta do preferowania rozwiązań proceduralnie bezpiecznych. Jeśli eksperyment się powiedzie, sukces przypisze się wielu czynnikom; jeśli zawiedzie, dokument z podpisem konkretnego urzędnika precyzyjnie wyznaczy odpowiedzialność. Strukturalnie wzmacnia to status quo bias.

Właśnie dlatego postulat akceptacji kontrolowanego ryzyka według Stevenson-Dodda jest w sektorze publicznym trudniejszy do realizacji niż w przedsiębiorstwie, gdyż publiczne pieniądze podlegają szczególnej odpowiedzialności. Nie można tu po prostu zaimportować start-upowego „fail fast". Potrzebny byłby formalnie legitymizowany reżim pilotażu: ograniczona skala, jawna hipoteza, kryteria zatrzymania, standard etyczny, akceptowany poziom ryzyka oraz określone zasady rozliczenia wiedzy z testu. Innymi słowy, sektor publiczny potrzebuje legalnej przestrzeni do inteligentnej porażki, a nie nieformalnej tolerancji niepowodzeń. Problemem pozostaje również niezgodność rytmów czasowych: budżet i sprawozdania mają cykl roczny, podczas gdy znaczna część wpływu społecznego – w edukacji, profilaktyce zdrowotnej czy reintegracji – generuje skutki dopiero w perspektywie wieloletniej.

Bariery czasowe i fragmentacja danych w pomiarze wpływu

Roczny system finansowy z natury preferuje wskaźniki, które można wykazać w jednym okresie rozliczeniowym. Prowadzi to do zjawiska temporal displacement: długookresowy outcome zostaje zastąpiony miernikiem krótkoterminowym, gdyż tylko ten pozwala na rozliczenie w bieżącej strukturze. Rozwiązaniem nie jest likwidacja rocznego budżetu, lecz wprowadzenie warstw czasowych. Roczne outputs mogą pozostać narzędziem kontroli, ale powinny być osadzone w wieloletniej Theory of Change. Program musi precyzyjnie określać, które rezultaty mierzy dzisiaj,

które za dwa lata, a które dopiero po zakończeniu finansowania. Bez tego wskaźnik przestaje reprezentować cel i zaczyna reprezentować kalendarz księgowy.

Polska administracja dysponuje ogromnymi zasobami danych w systemach podatkowych, ubezpieczeniowych, zdrowotnych, edukacyjnych, na rynku pracy, w pomocy społecznej oraz statystyce publicznej. Problemem nie jest zatem sam niedobór informacji, lecz ich fragmentacja instytucjonalna, prawna i semantyczna. Jedna instytucja wie, że uczestnik ukończył program, inna – że podjął zatrudnienie, a jeszcze kolejna – że po kilku miesiącach ponownie stał się bezrobotny.

Jeżeli dane te nie zostaną połączone w ramach legalnego i bezpiecznego procesu badawczego, każda z instytucji widzi jedynie fragment ścieżki. Z perspektywy Stevensona-Dodda jest to paradoks szczególnie istotny: państwo może posiadać więcej danych niż kiedykolwiek i nadal nie widzieć pełnego outcome. Rozwiązanie nie polega jednak na bezkrytycznym łączeniu baz. Ochrona danych, ograniczenie celu przetwarzania oraz zasada minimalizacji są kluczowe. Potrzebne są raczej kontrolowane środowiska badawcze, bezpieczna pseudonimizacja, wspólne identyfikatory oraz interoperacyjność semantyczna. Państwo musi nauczyć się wykorzystywać dane do ewaluacji tak, by impact measurement nie stał się systemem nadzoru nad obywatelem.

GUS w ostatnich latach rozwija eksperymentalne dane dotyczące gospodarki społecznej, w tym udziału podmiotów ekonomii społecznej w świadczeniu usług społecznych w latach 2022–2024. To istotny przykład poszerzania infrastruktury wiedzy o sektorze. Należy jednak konsekwentnie utrzymywać rozróżnienie między statystyką stanu a ewaluacją przyczynową. GUS może wykazać, ile podmiotów działa, kogo zatrudniają, w jakich regionach operują i jak zmieniają się te wartości w czasie. Nie oznacza to jednak automatycznie, że potrafimy przypisać zaobserwowaną zmianę konkretnej polityce. Statystyka publiczna tworzy mapę rzeczywistości. Ewaluacja odpowiada na pytanie, co zmieniło mapę. Oba systemy są sobie niezbędne, lecz pozostają epistemicznie różne.

Polska jest również państwem o dużym zróżnicowaniu regionalnym i lokalnym. Ten sam wskaźnik może mieć odmienne znaczenie w metropolii, mieście średnim, powiecie peryferyjnym czy gminie wiejskiej. Wskaźnik zatrudnienia nabiera innej interpretacji tam, gdzie funkcjonuje rozwinięty transport publiczny, i tam, gdzie brak samochodu sam w sobie stanowi barierę w dostępie do pracy.

Layering i asymetria między kontrolą a uczeniem

Centralny system wymaga porównywalności, jednak lokalna polityka potrzebuje adekwatności. To klasyczny konflikt między standardisation a contextualisation. Dlatego polski impact management powinien opierać się na architekturze "common core + local layer": niewielki zestaw wspólnych

outcomes umożliwiłby porównanie i agregację danych, podczas gdy region lub gmina zachowałoby dodatkowe wskaźniki wynikające z lokalnej Theory of Change. Takie rozwiązanie jest zgodne z zasadą, że złożoność powinna pozostawać w warstwie analitycznej, a nie obciążać każdego decydenta tym samym katalogiem kilkudziesięciu miar.

Wyjaśniając polski formalizm, należy unikać kulturowej esencjalizacji. Nie ma naukowych podstaw, by przypisywać Polakom wrodzoną skłonność do biurokracji; znacznie bardziej użyteczne jest podejście historyczno-instytucjonalne. Współczesny system powstał poprzez nakładanie się kolejnych warstw: klasycznej administracji legalistycznej, transformacji ustrojowej, decentralizacji, integracji europejskiej, polityki spójności, nowego zarządzania publicznego, cyfryzacji oraz nowych reżimów compliance. Każda reforma dodawała mechanizmy niezbędne z własnego punktu widzenia. W rezultacie instytucja może jednocześnie raportować według klasyfikacji budżetowej, zadaniowej, programu europejskiego, strategii sektorowej i systemu kontroli wewnętrznej. To właśnie layering – nakładanie kolejnych warstw bez pełnej eliminacji poprzednich – wyjaśnia współczesne przeciążenie pomiarowe lepiej niż samo hasło „biurokracja". Nowy wskaźnik jest bowiem znacznie łatwiej dodać, niż stary usunąć.

Polski system podlega zatem zjawisku, które można określić jako asymetrię rozwoju accountability i learning. Kontrola, sprawozdawczość i dokumentacja rozwijają się wraz ze wzrostem skali środków oraz wymogów regulacyjnych, jednak zdolność do interpretowania danych i modyfikowania polityki nie zawsze rośnie w tym samym tempie. Głównym deficytem nie jest więc brak kolejnych mierników, lecz brak institutional slack – czasu, kompetencji i przestrzeni niezbędnej do refleksji. Urzędnik obsługujący ogromną liczbę procedur może dysponować wszystkimi danymi potrzebnymi do uczenia się organizacji, nie mając jednocześnie czasu, by zadać im właściwe pytanie.

To kolejny powód, dla którego minimalizm pomiarowy w ujęciu Stevensona i Dodda jest w Polsce atrakcyjny. Nie chodzi o to, że państwo powinno zbierać mało informacji, lecz o to, że warstwa decyzyjna powinna być radykalnie mniej przeciążona niż warstwa danych. Z tego obrazu wyłania się paradoksalny wniosek: Polska nie cierpi przede wszystkim na brak raportów, badań, wskaźników czy instytucji ewaluacyjnych.

Od rytualnej ewaluacji do państwa uczącego się

W wielu obszarach wskaźników jest bardzo dużo. W niektórych – zwłaszcza w polityce spójności – dysponujemy nawet dojrzałymi metodami wnioskowania o wpływie. Dlatego postulat, by „mierzyć więcej", może jedynie pogłębić problem. Bardziej adekwatny program reformy powinien opierać się

na redukcji redundancji, zwiększeniu interoperacyjności danych, wyraźnym rozdzieleniu informacji służącej kontroli od tej służącej uczeniu oraz silniejszym powiązaniu rezultatów (outcome) z przyszłą alokacją zasobów. Organizacja publiczna nie powinna utrzymywać pięćdziesięciu strategicznych KPI tylko dlatego, że potrafi je wygenerować; powinna dysponować kilkoma miernikami, które realnie zmieniają decyzje, oraz bogatszym zapleczem analitycznym pozwalającym je interpretować.

Pytanie w duchu Stevensona-Dodda można zatem przekształcić w konkretny test dla polskiej instytucji. Nie należy zaczynać od pytania o posiadanie strategii, ewaluacji czy dashboardu, lecz ustalić, jaka decyzja została zmieniona przez informację o rezultacie. Czy program o słabym outcome został przeprojektowany? Czy przesunięto środki? Czy po odkryciu nierówności zmieniono kryteria dostępu? Czy zakończono interwencję mimo prawidłowego wykonania budżetu? Czy rozszerzono program, którego impact został wiarygodnie wykazany? Czy w wyniku ewaluacji zmieniono regulamin kolejnego konkursu? Czy system rejestruje przypadki, w których rekomendacja nie została wdrożona, wraz z odpowiednim uzasadnieniem? Dopiero odpowiedzi na te pytania pozwalają odróżnić państwo posiadające ewaluację od państwa uczącego się poprzez ewaluację.

Najbardziej ambitnym kierunkiem byłoby rozwinięcie polskiego systemu o poziom określany jako audit Theory of Change. Nie chodziłoby tu o zastąpienie istniejących kontroli, lecz o wprowadzenie pytania uprzedniego: na jakich dowodach opierała się teoria, że finansowane działanie doprowadzi do zamierzonego rezultatu? W takim systemie już przed startem programu należałoby ujawnić kluczowe założenia, alternatywne mechanizmy i warunki niepowodzenia. Dzięki temu ewaluacja nie musiałaby rekonstruować modelu logicznego po fakcie, lecz otrzymałaby hipotezę sformułowaną ex ante. Pozwoliłoby to również odróżnić implementation failure od theory failure. Program mógłby bowiem nie osiągnąć rezultatu z powodu błędnego wdrożenia lub dlatego, że jego podstawowa teoria była wadliwa – obie sytuacje wymagają zupełnie innej reakcji.

Jednym z kluczowych wniosków dla Polski jest potrzeba rozdzielenia dwóch ścieżek systemu informacyjnego. Accountability track powinien odpowiadać za legalność, standardy, wykonanie finansowe, ochronę praw, uczciwość oraz stabilne i audytowalne rezultaty. Learning track natomiast powinien umożliwiać testowanie hipotez, analizę przypadków negatywnych, wykorzystanie danych jakościowych, przeprowadzanie pilotaży i modyfikację programów. Jeśli każda informacja generowana na potrzeby uczenia zostanie automatycznie użyta do sankcjonowania wykonawcy, proces uczenia zostanie zahamowany. Jeśli zaś warstwa uczenia nie będzie miała związku z odpowiedzialnością, łatwo przeobrazi się w miękką narrację o

„wyciągniętych lekcjach". Konieczna jest zatem kontrolowana separacja, a następnie sprzężenie: swobodne generowanie wiedzy i formalny moment, w którym staje się ona podstawą decyzji.

Ważny jest również wniosek aksjologiczny. Krytyka polskiego legalizmu nie może przerodzić się w naiwną pochwałę menedżerskiej swobody. W państwie demokratycznym procedura chroni wartości, których pojedynczy outcome nie mierzy: równość, przewidywalność, możliwość odwołania, ochronę mniejszości i ograniczenie władzy dyskrecjonalnej.

Polski wariant „Difference" nie powinien więc imitować przedsiębiorstwa, lecz budować państwo zdolne do uczenia się bez utraty państwa prawa. Zadanie to jest trudniejsze niż zwykła deregulacja, gdyż wymaga pogodzenia dwóch racjonalności często uznawanych za sprzeczne: formalnej odpowiedzialności i adaptacyjnego zarządzania. W rzeczywistości dopiero ich synteza tworzy dojrzałą instytucję publiczną. Gdyby do polskiego systemu po prostu dodano obowiązek sporządzania „raportu impact", prawdopodobnym rezultatem byłaby kolejna warstwa compliance. Byłaby to ironiczna realizacja problemu, który Stevenson-Dodd stara się rozwiązać: nowy formularz mający zwalczać nadmiar formularzy.

Przejście od rytuału wykonania ku kulturze uczenia się

Potrzebna jest zatem nie kolejna kategoria dokumentu, lecz infrastruktura: kompatybilne definicje outcomes, bezpieczny dostęp do danych administracyjnych, wspólne narzędzia psychometryczne (tam, gdzie mają zastosowanie), repozytorium Theory of Change, system śledzenia rekomendacji, finansowanie follow-up, standardy contribution analysis oraz kompetencje analityczne po stronie instytucji finansujących i wykonawczych. Infrastruktura różni się od obowiązku tym, że obniża koszt dobrego pomiaru. Gdyby każda mała organizacja musiała samodzielnie budować metodologię, narzędzia IT, bazy wskaźników i system longitudinal tracking, pomiar stałby się przywilejem dużych podmiotów. Wspólna infrastruktura może zatem demokratyzować zdolność do uczenia się.

Po zestawieniu danych trudno obronić zarówno narrację pesymistyczną, jak i triumfalistyczną. Polska nie jest systemem, który „tylko liczy procedury". Posiada budżet zadaniowy, rozbudowaną ewaluację polityki spójności, system rekomendacji, rozwijające się programy ekonomii społecznej, instrumenty finansowania instytucjonalnego oraz prawo zamówień dopuszczające kryteria społeczne. Jednocześnie krytyka NIK dotycząca ograniczonej sprawczości budżetu zadaniowego, wysoki poziom obciążeń biurokratycznych deklarowany przez NGO, problemy z płynnością sektora oraz strukturalna projektowość uzasadniają hipotezę, że w wielu obszarach aparat mierzenia rozwijał się szybciej niż aparat wykorzystywania pomiaru do adaptacji. Dlatego najtrafniejszą

kategorią diagnostyczną nie jest „Polska jako organizacja Do", lecz Polska jako system hybrydowy znajdujący się pomiędzy kulturą zgodności, kulturą ewaluacji a dopiero wyłaniającą się kulturą zarządzania wpływem.

Pierwsza z nich pyta: czy działanie było zgodne z prawem i planem. Druga: czy przyniosło rezultat. Trzecia: co powinniśmy zrobić inaczej po poznaniu tego rezultatu. W polskim systemie wszystkie trzy racjonalności są już obecne, lecz nie wszędzie dysponują jednakową siłą instytucjonalną. Największy potencjał koncepcji Stevensona-Dodda polega więc w Polsce nie na imporcie nowej metodologii pomiarowej, lecz na przesunięciu środka ciężkości istniejącej infrastruktury. Państwo i sektor społeczny dysponują przecież wieloma danymi, miernikami i badaniami.

Kolejnym etapem powinno być skracanie drogi od informacji do decyzji, od rekomendacji do zmiany instrumentu, od wykrycia nieskuteczności do dezinwestycji oraz od rozpoznania sukcesu do odpowiedzialnego skalowania. Przejście to można opisać jako ewolucję od accountability for spending, przez accountability for results, ku accountability for learning. Nie wystarczy bowiem rozliczyć instytucję z tego, czy wydała pieniądze zgodnie z prawem, ani nawet czy osiągnęła zadeklarowany wynik. Dojrzała organizacja publiczna powinna umieć wykazać, czego nauczyła się o mechanizmie działania, które założenie odrzuciła, jak zmieniła kolejną decyzję oraz dlaczego nadal finansuje to, co finansuje. W tym punkcie analiza powraca do rdzenia modelu Stevensona-Dodda.

Różnica między „Do" a „Difference" nie sprowadza się ostatecznie do liczby wskaźników czy nowoczesności dashboardu. Polega na zdolności instytucji do stwierdzenia: „wykonaliśmy zadanie, ale nie osiągnęliśmy wystarczającej zmiany, więc zmieniamy sposób działania". Dopiero taka deklaracja świadczy, że rezultat uzyskał pierwszeństwo przed rytuałem wykonania. W takim systemie ewaluacja przestaje być jedynie komentarzem do przeszłości, a staje się częścią technologii rządzenia przyszłością.

> Prawdziwa dojrzałość instytucjonalna nie objawia się w perfekcyjnie wypełnionym raporcie z wpływu, lecz w zdolności do przyznania: 'zrobiliśmy wszystko zgodnie z procedurą, ale świat się nie zmienił, więc musimy przestać to robić w ten sposób'. Czy jesteśmy zatem gotowi zamienić bezpieczny rytuał dokumentacji na ryzykowny proces autentycznego uczenia się? W systemie, gdzie błąd jest sankcjonowany, a poprawność formularza nagradzana, taka zmiana wymaga czegoś więcej niż nowej metodologii – wymaga redefinicji pojęcia odpowiedzialności.

Inspiracje

[1] 

"The Difference How to Understand Measure and Maximise Social Impact" Matt Stevenson-Dodd

Practical Inspiration Publishing | ISBN: 9781788609562

Matt Stevenson-Dodd jest brytyjskim przedsiębiorcą społecznym, dyrektorem organizacji non-profit i konsultantem specjalizującym się w pomiarze i raportowaniu wpływu społecznego. Ukończył z wyróżnieniem studia MBA na Uniwersytecie w Nottingham oraz program Strategic Perspectives in Nonprofit Management w Harvard Business School. W latach 2010-2019 Stevenson-Dodd pełnił funkcję dyrektora generalnego Street League, brytyjskiej organizacji charytatywnej wykorzystującej sport do walki z bezrobociem wśród młodzieży, gdzie był pionierem w zakresie transparentnego raportowania wpływu i otwartego raportowania wyzwań organizacyjnych. W 2019 roku założył Trust Impact, firmę konsultingową doradzającą organizacjom non-profit, firmom i instytucjom publicznym w zakresie oceny i maksymalizacji wpływu społecznego. Pełnił również funkcję przewodniczącego Football Beyond Borders oraz członka zarządu kilku organizacji charytatywnych, w tym St John Ambulance, za którą w 2020 roku został mianowany członkiem Zakonu św. Jana (MStJ).

Matt Stevenson-Dodd *is a British social entrepreneur, non-profit executive, and consultant specializing in social impact measurement and reporting. He holds an MBA with distinction from the University of Nottingham and completed Harvard Business School's Strategic Perspectives in Nonprofit Management program. From 2010 to 2019, Stevenson-Dodd served as Chief Executive Officer of Street League, a UK charity using sports to tackle youth unemployment, where he pioneered transparent impact reporting and open reporting of organizational challenges. In 2019, he founded Trust Impact, a consultancy advising non-profit organizations, businesses, and public bodies on evaluating and maximizing social impact. He has also served as chair of Football Beyond Borders and as a trustee for several charitable bodies, including St John Ambulance, for which he was appointed a Member of the Order of St John (MStJ) in 2020.*

Trust Impact

Literatura uzupełniająca

Książki

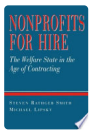
Książki autorów przywoływanych i literatura pokrewna:



[1] "The Risk Management of Everything"

Michael Power

2004 | Demos | ISBN: 9781841801278



[2] "Nonprofits for Hire: the Welfare State in the Age of Contracting"

Michael Lipsky

Harvard University Press | ISBN: 9780674043817

[3] "Toward a Theory of Street-level Bureaucracy"

Michael Lipsky

1969

Artykuły naukowe

Prace autorów przywoływanych w tekście:

[1] "The Audit Society: Rituals of Verification"

Michael Power

1999 | RePEc: Research Papers in Economics

Źródło

[2] "The risk management of everything"

Michael Power

2004 | The Journal of Risk Finance | DOI: 10.1108/eb023001

Źródło

[3] "Accounting, Organizing, and Economizing: Connecting Accounting Research and Organization Theory"

Peter Miller, Michael Power

2013 | Academy of Management Annals | DOI: 10.1080/19416520.2013.783668

Źródło

[4] "The risk management of nothing"

Michael Power

2009 | Accounting Organizations and Society | DOI: 10.1016/j.aos.2009.06.001

Źródło

Treść tworzy Fundacja Dobre Państwo.
Redaguje i publikuje APA ONE, autorski system redakcyjny Fundacji oparty na AI.
APA ONE Publishing System
Fundacja Dobre Państwo © 2026
Article ID: 2a97a575-96ac-4497-a31b-62f94ade1c3d