

## Ramy instytucjonalne, korzyści, wyzwania

# Big data i cyfrowa transformacja w NOK

Ostatnie lata przyniosły szybki rozwój technologii przetwarzania danych, poczynając od uczenia maszynowego, przez big data (zbiory danych) do sztucznej inteligencji (ang. *artificial intelligence*, AI). W konsekwencji zwiększają się możliwości ich wykorzystania we wszystkich dziedzinach wiedzy (m.in. bezpieczeństwo, finanse, handel, marketing). Komisja Europejska oszacowała udział gospodarki opartej na danych w unijnym PKB na 3,9%<sup>1</sup> w 2022 r. Obszarem wykorzystania big data stopniowo staje się działalność najwyższych organów kontroli (NOK). Wiąże się to jednak z wieloma wyzwaniami oraz koniecznością dokonania zmian metodycznych i organizacyjnych, ale może doprowadzić do skuteczniejszej kontroli administracji publicznej i innych podmiotów, a jednocześnie do lepszego wykorzystania zasobów<sup>2</sup>.

### MARCIN BĘDZIESZAK

#### Cel i zakres artykułu

Celem artykułu jest przedstawienie możliwości wykorzystania narzędzi big data w działalności najwyższych organów kontroli. W pierwszej części przybliżono znaczenie

tej technologii w sektorze publicznym, w drugiej nakreślono uwarunkowania związane z wykorzystaniem wielkich zbiorów danych, a w ostatniej zaprezentowano rozważania dotyczące transformacji cyfrowej NOK i czynniki mające wpływ na wykorzystanie narzędzi big data w tych organach.

<sup>1</sup> *European Data Market Study 2021 2023, D2.4 Second report on facts and figures*, IDC-Lison Research Council, European Commission, luty 2023.

<sup>2</sup> M. Maciejewski: *Skuteczność i efektywność kontroli administracji – zbiory danych (big data) i możliwości ich wykorzystania*, „Kontrola Państwowa” nr 5/2015.

## Koncepcja big data i jej znaczenie

Pod pojęciem big data rozumie się duże ilości danych do analizy, których liczbę należy maksymalizować, aby uzyskać największą wartość informacyjną<sup>3</sup>. Może też odnosić się ono do nowych rozwiązań technologicznych, dotyczących przetwarzania wielkich wolumenów danych o całkowicie innym charakterze (ilościowym i jakościowym) niż dotychczas<sup>4</sup>. Big data jest tradycyjnie definiowane przez 3V, tj.:

- przetwarzanie dużych wolumenów danych (ang. *volume*);
- dużą zmienność i dynamikę przetwarzanych danych (ang. *velocity*), w szczególności w czasie zbliżonym do rzeczywistego (np. danych sensorycznych, strumieniowych, pochodzących z Internetu);
- dużą różnorodność danych (ang. *variety*), zwłaszcza danych nieustrukturyzowanych<sup>5</sup>.

Pozostałe „v” dodawane do klasycznego modelu dotyczą:

- wartości danych (ang. *value*) na potrzeby podejmowania decyzji w organizacji;
- wiarygodności danych (ang. *veracity*), w tym problemów ich jakości wynikających z przetwarzania danych pochodzących z różnorodnych źródeł.

Big data jest uważana za „nową ropę naftową” sztucznej inteligencji<sup>6</sup>. Główną zaletą korzystania z narzędzi analityki danych jest możliwość przetwarzania ogromnych ilości danych w krótkim czasie, co pozwala użytkownikom na szybsze i dokładniejsze wykrywanie anomalii, prognozowanie i/lub znajdowanie prawidłowości w dużych zbiorach. Najważniejszymi dziedzinami, w których wykorzystuje się technologie analityki danych są telekomunikacja, transport oraz nauki przyrodnicze i medyczne. W telekomunikacji główne obszary to sieci komputerowe i Internet, radio i telewizja, nadawanie, telefonia, wideokonferencje i cyfrowe centrale telefoniczne (VoIP). Zastosowania w transporcie obejmują lotnictwo i awionikę, pojazdy autonomiczne, rozpoznawanie kierowcy/pojazdu, transport i inżynierię ruchu<sup>7</sup>. W naukach przyrodniczych i medycznych są one przydatne w bioinformatyce, inżynierii biologicznej, biomechanice, badaniach nad lekami, genetyce i genomice, obrazowaniu medycznym, neuronauce i neurorobotyce, informatyce medycznej, żywieniu i nauce o żywności, monitorowaniu parametrów fizjologicznych i zdrowiu publicznym. Oprócz tego technologie big data i sztuczna inteligencja

<sup>3</sup> M. Cox, D. Ellsworth: *Managing Big Data for Scientific Visualization*, Exploring Gigabyte Datasets in Real-Time: Algorithms, Data Management, and Time-Critical Design, Los Angeles 1997.

<sup>4</sup> *Big data: The next frontier for innovation, competition, and productivity*, McKinsey Global Institute, 2011, <<https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/big-data-the-next-frontier-for-innovation>> (dostęp 14.11.2023).

<sup>5</sup> J. Wiczorkowski: *Wykorzystanie koncepcji big data w administracji publicznej*, „Roczniki Kolegium Analiz Ekonomicznych” nr 33/2014.

<sup>6</sup> A. Agrawal, J. Gans, A. Goldfarb: *Prediction machines: the simple economics of artificial intelligence*, Harvard Business Press Boston, MA 2018.

<sup>7</sup> *Use of big data and AI in fighting corruption and misuse of public funds. Good practice, ways forward and how to integrate new technology into contemporary control framework, organized by the Policy Department on Budgetary Affairs for the Committee on Budgetary Control*, European Parliament 2021.

są stosowane w dziedzinie finansów, m.in. do wykrywania oszustw i prania pieniędzy.

Szczególny obszar wykorzystania technologii big data to sektor publiczny i administracja publiczna, w których jest konieczne podejście systemowe<sup>8</sup>, a w badaniach wskazuje się, że kluczowymi czynnikami sukcesu we wdrażaniu tych rozwiązań są: organizacja, ludzie, technologia, struktura zarządzania i zarządzanie danymi<sup>9</sup>.

Wykorzystanie analityki danych w sektorze publicznym wynika z jednej strony z rozwoju technologicznego, takiego jak zwiększona dostępność danych do ponownego wykorzystania oraz rozwój technologii ich przetwarzania. Nie mniej istotne z punktu widzenia projektowania jest jednak źródło pozyskania danych, które będą wykorzystywane na potrzeby administracji publicznej. Możliwe jest wyszczególnienie dwóch głównych rodzajów danych sektora publicznego i ich zastosowania:

- dane dotyczące osób korzystających z usług oferowanych przez administrację publiczną mogą być wykorzystane do lepszego rozumienia potrzeb mieszkańców, określenia jakości oferowanych usług oraz wskazania zmian w prawie oraz procesach

realizowanych przez jednostki administracji publicznej;

- dane generowane przez jednostki administracji publicznej mogą być wykorzystane do poprawy procesów, jakie w nich zachodzą, budowy mechanizmów bezpieczeństwa opartych na profilowaniu działań pracowników, ułatwienia dostępu do danych uczestników procesów, poprawy transparentności działań podejmowanych przez jednostki tego sektora<sup>10</sup>.

Dostępność zwiększa się dzięki publikacji i ponownemu wykorzystaniu danych rządowych jako zbiorów otwartych oraz dzięki połączeniu rozmaitych zestawów danych z różnych organizacji (jako dużych zbiorów danych). Z drugiej strony analitykę sektora publicznego kreują organizacje rządowe, których celem jest rozwiązywanie problemów społecznych i informowanie o różnych etapach tworzenia polityki dzięki opieraniu się w większym stopniu na dowodach.

W konsekwencji spektrum zastosowania big data w sektorze publicznym obejmuje:

- proces legislacyjny i tworzenie polityki publicznej<sup>11</sup>,
- usługi publiczne dla obywateli i biznesu<sup>12</sup>,

<sup>8</sup> Przykład takiego podejścia prezentuje OECD w opracowaniu *The Path to Becoming a Data-Driven Public Sector*, OECD Digital Government Studies, OECD Publishing, Paris 2019. Podkreśla się w nim, że zarządzanie danymi w sektorze publicznymi powinno być realizowane w cyklu: przywództwo i wizja (np. polityka i strategia danych) → zdolność do spójnego wdrażania (np. umiejętności i szkolenia, finansowanie, eksperymenty i innowacje w zakresie danych) → regulacje (zasady, wytyczne, przewodniki) → cykl wartości danych (zarządzanie danymi) → infrastruktura danych (rejstry, katalogi, jeziora danych, rozwiązania chmurowe) → architektura dostarczania.

<sup>9</sup> W.R. Titera: *Updating Audit Standard--Enabling Audit Data Analysis*, "Journal of Information Systems" nr 1/2013.

<sup>10</sup> M. Pańkowska, M. Żytniewski: *Modele architektury biznesowej administracji publicznej w warunkach przetwarzania danych masowych*, „Roczniki Kolegium Analiz Ekonomicznych nr 56/2019.

<sup>11</sup> Analityka predykcyjna i nowe źródła danych, w tym pełniejsze i szybciej dostępne statystyki publiczne, mogą istotnie wspierać proces legislacyjny i politykę publiczną prowadzoną przez państwo.

<sup>12</sup> Potencjalne korzyści obejmują poprawę jakości i skuteczności oferowanych usług publicznych, ich indywidualizację, rozumianą jako dopasowanie do potrzeb konkretnych odbiorców i nagłych sytuacji oraz wskazanie nowych obszarów, w których realizacja usług publicznych byłaby pożądana.

- bezpieczeństwo państwa i walkę z przestępstwami<sup>13</sup>,
- statystykę publiczną<sup>14</sup>.

W raporcie TechAmerica Foundation wymieniono dalsze, czasami bardziej szczegółowe obszary zastosowania big data przez urzędy administracji publicznej: poprawa jakości i efektywności funkcjonowania służby zdrowia, monitorowanie istotnych danych w celu wczesnego wykrywania zagrożeń epidemiologicznych i sanitarnych, wspomaganie zarządzania transportem, ocena i poprawa jakości systemu edukacji, wykrywanie nieprawidłowości w finansowaniu służby zdrowia, zapewnienie bezpieczeństwa w cyberprzestrzeni, wykrywanie nieprawidłowości w rozliczeniach podatkowych, lepsze prognozowanie pogody i przewidywanie klęsk żywiołowych, analiza rynku pracy i przeciwdziałanie bezrobociu<sup>15</sup>.

### Wykorzystanie big data w sektorze publicznym

Do skutecznego wykorzystania potencjału big data w sektorze publicznym, w tym w najwyższych organach kontroli niezbędne jest przygotowanie i wdrożenie rozwiązań

instytucjonalnych, charakteryzujących się elastycznością, a także innowacyjnych i dostosowanych do potrzeb administracji publicznej. Jak wskazano wcześniej, jednym z wyzwań w wykorzystaniu big data w administracji publicznej jest liczebność podmiotów, co w konsekwencji powoduje rozproszenie źródeł danych. Pierwszym elementem powinno być zatem otwieranie, łączenie i udostępnianie danych rządowych między nimi. Drugim elementem, podobnie jak w wypadku innych reform sektora publicznego, stanowiącym kamień węgielny realizacji projektów danych, jest zaangażowanie ze strony kierownictwa oraz przywództwo. Niezbędne jest, aby kierownictwo i kluczowi decydenci ustalili jasną, kompleksową wizję wykorzystania danych, która mieści się w szerszym planie rozwoju i obejmuje dostępne procedury oraz zachęty dla twórców, analizatorów i użytkowników danych. To wewnętrzne zaangażowanie jest czynnikiem zapewniającym:

- dostęp do odpowiednich i aktualnych danych;
- przemiany kulturowe potrzebne w organizacjach, aby doprowadzić do procesów decyzyjnych opartych na danych<sup>16</sup>.

<sup>13</sup> Dzięki wielkim zbiorom danych państwo zyskuje dodatkowe narzędzia kontroli i monitorowania aktywności jednostek, grup, przedsiębiorstw oraz innych organizacji, co może przekładać się na skuteczniejsze działania prewencyjne oraz na skuteczniejszą walkę z przestępstwami. K. Kosior: *Big data w sektorze publicznym – szanse, ograniczenia, perspektywy*, „Kultura i Polityka” nr 20/2016.

<sup>14</sup> Wykorzystanie danych transakcyjnych z dużych sieci handlowych (*scanner data*) umożliwiłoby dokładniejszy pomiar cen oraz inflacji; z kolei danych pochodzących od dostawców telefonii komórkowych – możliwości wykorzystania ich w zagadnieniach związanych z mobilnością ludzi, przepływami do pracy czy danych pochodzącymi ze specjalistycznych portali internetowych. M. Beręsewicz, M. Szymkowiak: *Big data w statystyce publicznej – nadzieje, osiągnięcia, wyzwania i zagrożenia*, „Ekonometria = Econometrics” nr 2/2015.

<sup>15</sup> J. Wieczorkowski: *Wykorzystanie koncepcji big data w administracji publicznej*, „Roczniki Kolegium Analiz Ekonomicznych” nr 33/2014.

<sup>16</sup> L. Tomar, W. Guicheney, H. Kyarisiima, T. Zimani, B. Roseth, S. Acevedo: *Big Data in the Public Sector. Selected Applications and Lessons Learned*, Institutions for Development Sector – Institutional Capacity of the State Division, Discussion Paper N° IdB-Dp-483, Inter-American Development Bank, październik 2016.

Należy także założyć, na co będą wskazywać przykłady zastosowania big data, że sposób ich wykorzystania jest projektowany na potrzeby danej agencji publicznej.

Wymaga to inwestycji w narzędzia techniczne, takie jak chmura lub hurtownie danych, gdzie informacje mogą być przechowywane i wykorzystywane przez twórców i odbiorców. Często zachodzi potrzeba określenia przepisów i praktyk dotyczących udostępniania już przechowywanych danych rządowych.

Ostatnim, choć nie najmniej ważnym elementem jest kapitał ludzki wymagany do realizacji zadań technicznych i analitycznych. Pracownicy powinni być zaznajomieni ze strategiczną wizją wykorzystania danych i mieć doświadczenie w danologii (ang. *data science*), a także w wykorzystywaniu konkretnych narzędzi. Konieczne są też procedury dotyczące sposobu postępowania z danymi oraz szkolenia urzędników służby cywilnej na temat przydatności big data w realizacji codziennych zadań.

W literaturze wskazuje się następujące wyzwania dotyczące big data w sektorze publicznym:

- zapewnienie jakości i bezpieczeństwa danych;
- konieczność zadbania o infrastrukturę, która zapewni ponowne ich wykorzystanie (interoperacyjność);
- ograniczona gotowość organizacyjna sektora publicznego;

- niechęć rządów do polegania na nowych lub powstających źródłach danych;
  - niechęć decydentów do bycia informowanym przez *data science*;
  - negatywne konsekwencje dla pracowników operacyjnych;
  - pojawienie się stronniczości w algorytmach;
  - kwestie prawne obejmujące prywatność i odpowiedzialność rządowego procesu decyzyjnego<sup>17</sup>;
- a dodatkowo:

- brak wykwalifikowanych pracowników zorientowanych biznesowo, świadomych, gdzie i jak duże zbiory danych mogą pomóc w rozwiązywaniu problemów sektora publicznego;
- duża liczba organów administracji publicznej (zwłaszcza zdecentralizowanych)<sup>18</sup>.

Wykorzystanie big data, w tym przez instytucje publiczne, niesie ryzyko inwigilacji społeczeństwa. Systemy rozpoznawania twarzy stosowane w dochodzeniach przez część krajów Unii Europejskiej, dane z monitoringu drogowego, dane telekomunikacyjne, transakcje płatnicze są wykorzystywane w ściśle określonych celach, ale jak pokazuje przykład chińskiego systemu zaufania społecznego<sup>19</sup>, mogą także służyć tzw. inżynierii społecznej. Kluczowymi kwestiami pozostają zatem granice wykorzystania technologii big data przez władze publiczne oraz aspekt etyczny tych działań.

<sup>17</sup> A.F. van Veenstra, F. Grommé, S. Djafari: *The use of public sector data analytics in the Netherlands*, "Transforming Government: People, Process and Policy" nr 4/2021.

<sup>18</sup> R. Munne: *Big Data in the Public Sector* [w:] *New Horizons for a Data-Driven Economy. A Roadmap for Usage and Exploitation of Big Data in Europe*, J.M. Cavanillas, E. Curry, W. Wahlster (red.), Springer 2016.

<sup>19</sup> M. Bartoszewicz: *Chiński system zaufania społecznego*, „Przegląd Geopolityczny” nr 32/2020.

Drugim istotnym elementem jest kwestia propagandy (obliczeniowej), rozumianej jako zbiór platform mediów społecznościowych, agentów autonomicznych i dużych baz danych, których zadaniem jest manipulowanie opinią publiczną<sup>20</sup>. Przykładem tego zjawiska są wybory prezydenckie w USA w 2016 r., w przypadku których była stosowana ukierunkowana propaganda polityczna i używano boty (programy komputerowe do zautomatyzowanych działań w sieci) do rozpowszechniania przekazów w mediach społecznościowych. Było to możliwe dzięki wielkim bazom danych, których używano do kształtowania opinii odbiorców oraz stworzeniu botów i algorytmów rozpowszechniających propagandę.

### Ramy wykorzystania big data w NOK

Zgodnie z zapisami Deklaracji z Limy w sprawie zasad kontroli finansów publicznych, „metody kontroli stosowane przez NOK powinny stale nadążać za postępem nauki i techniki w zakresie zarządzania finansami”<sup>21</sup>. Wobec zmian, jakie mają miejsce w analizie danych na Kongresie INTOSAI w 2019 r.<sup>22</sup> przygotowano dwie uchwały zachęcające organy kontroli

do stosowania nowoczesnych technologii i skutecznego reagowania na możliwości rozwoju stwarzane przez postęp technologiczny. To istotny krok w stronę modernizacji kontroli w sektorze publicznym, umożliwiającą kontrolerom wniesienie większej wartości w zarządzanie aktywami sektora publicznego, ale także w obszarze przejrzystości i odpowiedzialności<sup>23</sup>. W niektórych NOK analiza danych stała się integralną częścią procesu kontrolnego (np. Krajowy Urząd Kontroli Wielkiej Brytanii (National Audit Office, NAO)). Jego doświadczenia wskazują, że technologia ta pomaga kontrolerom:

- zwiększyć wydajność pracy dzięki automatyzacji rutynowych procedur i umożliwieniu przetwarzania dużych ilości danych;
- zyskać nową, pogłębianą wiedzę przez łączenie danych na temat anomalii i nieprzewidywalności;
- zachować spójność danych, unikać błędów lub opóźnień oraz przeprowadzać testy na całych populacjach, a nie dokonując doboru próby. Pozwala to przewyżnić ograniczenia tradycyjnie wykonywanych czynności kontrolnych<sup>24</sup>.

Najwyższe organy kontroli, jak wskazywało wcześniej, podążają za trendem wykorzystania narzędzi big data, o czym mogą

<sup>20</sup> S.C. Woolley, P.N. Howard: *Automation, Algorithms, and Politics | Political Communication, Computational Propaganda, and Autonomous Agents — Introduction*, "International Journal of Communication" nr 0 (October 12, 2016).

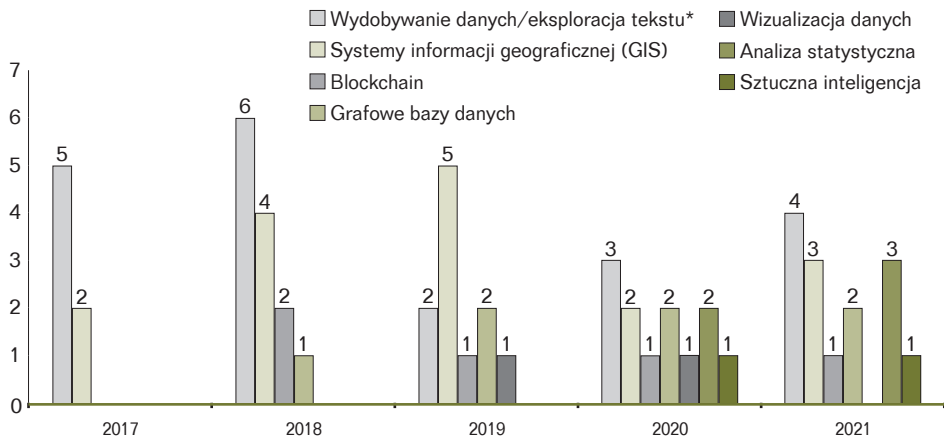
<sup>21</sup> Deklaracja z Limy w sprawie zasad kontroli finansów publicznych, <<https://www.nik.gov.pl/plik/id,15523.pdf>> (dostęp 18.8.2023).

<sup>22</sup> *The XXIII INCOSAI Outcomes Report*, INTOSAI, 2019.

<sup>23</sup> Z. Bozhinovska Lazarevska, T. Tocev, I. Dionisijev: *How to Improve Performance in Public Sector Auditing Through the Power of Big Data and Data Analytics? – The Case of The Republic of North Macedonia*, "Journal of Accounting, Finance and Auditing" Issue 3/2022.

<sup>24</sup> R. Kelly: *How data analytics can help with audits*, <<https://www.nao.org.uk/naoblog/how-data-analytics-can-help-with-audits/>> (dostęp 18.8.2023).

Rysunek 1. Liczba technologii zaprezentowanych na posiedzeniach Grupy roboczej INTOSAI ds. big data



\* Data mining/text mining

Źródło: Przegląd kontroli dotyczących wielkich zbiorów danych, przeprowadzonych przez najwyższe organy kontroli w latach 2016-2021, Grupa robocza INTOSAI ds. big data, wrzesień 2022.

świadczą, poza przykładami konkretnych kontroli, technologie wykorzystywane czy prezentowane podczas posiedzeń Grupy roboczej INTOSAI ds. big data (rys. I). Widoczne jest coraz większe różnicowanie oraz zaawansowanie technologii big data. Jeszcze kilka lat temu były to tradycyjne technologie, jak np. wydobywanie danych (ang. *data mining*), ale już obecnie są to zarówno te tradycyjne, jak i blockchain czy sztuczna inteligencja. Najwyższe organy kontroli Izraela i Indii planują kontrolę przygotowania do wykorzystania

sztucznej inteligencji w sektorze publicznym<sup>25</sup> oraz zastosowania jej w procesie kontroli<sup>26</sup>.

Badania<sup>27</sup> wskazują jednak, że większość NOK definiuje swoje podejście do wykorzystania technologii informacyjnych na podstawie własnych doświadczeń i ekspozycji (poziom zaawansowania technologicznego kraju) na potencjał transformacji technologicznej (innowacje). Jedynie nieliczne określiły strategię transformacji cyfrowej lub technologicznej z wykorzystaniem całościowej koncepcji zarządzania

<sup>25</sup> M. Kogosowski: *Israel State Comptroller to Audit Country's Readiness for AI Era*, <<https://www.israeldefense.co.il/en/node/58720>> (dostęp 18.8.2023).

<sup>26</sup> H. Sharma: *Look for ways to adopt AI into audit techniques for effectiveness: CAG*, <<https://indianexpress.com/article/business/economy/look-for-ways-to-adopt-ai-into-audit-techniques-for-effectiveness-cag/>> (dostęp 18.8.2023).

<sup>27</sup> J. E. Otia, E. Bracci: *Digital transformation and the public sector auditing: The SAI's perspective*, "Financial Accountability & Management" nr 38/2022.



zmianą. Warto też wskazać, że nie ma ogólnie przyjętych standardów regulujących integrację technologii w procesie kontroli. Można natomiast wymienić problemy związane z transformacją cyfrową, przed którymi stoją najwyższe organy kontroli:

- kontrolerzy nadal będą musieli stawiać czoła wielu wyzwaniom i możliwościom;
- transformacja cyfrowa wymaga od kontrolerów nowych zestawów umiejętności;
- środowisko informatyczne jednostki kontrolowanej może nie być dostosowane do wykorzystania analizy danych;
- struktury i formaty danych czasami wymagają przetworzenia i „oczyszczenia”, zanim zostaną uznane za dowody kontroli;
- koszt obsługi zautomatyzowanych systemów jest bardzo wysoki.

Naprzeciw tym problemom wychodzą działania Międzynarodowej Organizacji Najwyższych Organów Kontroli, (INTOSAI), która powołała trzy grupy robocze zajmujące się big data, a mianowicie Grupę roboczą ds. audytu IT, Grupę roboczą ds. dużych zbiorów danych oraz Grupę roboczą ds. wpływu nauki i technologii na audyt. Dodatkowo najwyższe organy kontroli na całym świecie prowadzą działania mające na celu budowanie potencjału (np. platforma e-learningowa IDI, akademia Europejskiego Trybunału Obrachunkowego), aby wzmocnić umiejętności cyfrowe kontrolerów i sprostać wyzwaniom, jakie niesie rewolucja cyfrowa. Zmianie ulegają także struktury organizacyjne NOK – np. tworzone są laboratoria innowacji,

zatrudniani analitycy danych, wykorzystywane jest środowisko tzw. chmur.

W literaturze wskazuje się, że proces transformacji cyfrowej NOK powinien być rozpatrywany przez pryzmat pięciu elementów tj. strategii, procesu, organizacji, ludzi i kultury oraz technologii. Badania w tym obszarze wskazują, że większość NOK daje pierwszeństwo wymiarowi technologicznemu i procesowemu<sup>28</sup>. Innym aspektem transformacji, takim jak kultura organizacyjna, przywództwo, ludzie i strategia poświęca się mniej uwagi. Większość organów kontroli zajmuje się bowiem obecnie automatyzacją procesów, wprowadza technologie do swoich starych systemów i tradycji kontrolnych, ale nadal brakuje zmian kulturowych związanych z metodyką kontroli czy rolą kontrolera.

Analityka danych, big data, a nawet sztuczna inteligencja znajdują zastosowanie w działalności audytowej, najbardziej zbliżonej do działalności NOK. Największą przewagą analityki danych nad prostymi procedurami analitycznymi jest jej zdolność do oceny ryzyka, identyfikacji anomalii i wykrywania błędów. W przeciwieństwie do tradycyjnych praktyk audytu, ten oparty na technologii zapewnia wyższą jakość dowodów, które pochodzą z wielu nowych źródeł, w tym dużych zbiorów danych oraz danych egzogenicznych, i są wynikiem analitycznego łączenia różnych procesów, potwierdzania baz danych i ciągłego monitorowania alertów<sup>29</sup>. Kierunki zmian skutkujące wykorzystaniem

---

<sup>28</sup> Tamże.

<sup>29</sup> W.R. Titera: *Updating Audit Standard--Enabling Audit Data Analysis*, "Journal of Information Systems" nr 1/2013.



analitiky danych w audycie (finansowym) wiążą się z jednej strony z szerszym wykorzystaniem tradycyjnych technik analizy danych, z użyciem takich narzędzi, jak Excel czy ACL lub technik rozszerzonych, np. analitiky predykcyjnej, a z drugiej z częstszym wykorzystaniem źródeł danych, w szczególności niefinansowych<sup>30</sup>. Ścieżka pierwsza – poszerzenie wykorzystywanego instrumentarium jest bliższa kontroli i audytowi finansowemu, z kolei druga – dotycząca źródeł danych – kontroli wykonania zadań.

Wykorzystanie narzędzi big data umożliwia kontrolerom i audytorom przetestowanie większej liczby transakcji, biorąc pod uwagę, że analiza danych zwiększa wystarczalność, czyli odpowiednią ilość dowodów badania. Model oparty na ryzyku i przykładowe transakcje są wykorzystywane do rzetelnego określenia zestawienia sald kont. Analitika danych umożliwia kontrolerom zautomatyzowanie testowania transakcji, bowiem teoretycznie cała populacja jest testowalna. Wskazuje się nawet, że w przyszłości przewidywane jest wykorzystanie zaawansowanej technologii, która umożliwi badanie wszystkich transakcji przez sortowanie, filtrowanie i analizowanie ogromnych ilości transakcji w celu wykrycia anomalii<sup>31</sup>. Taki rozwój działalności kontrolnej wymaga jednak intensywnej

współpracy środowisk akademickich, inżynierów, kontrolerów i wreszcie interesariuszy<sup>32</sup>. Niezbędne są dane i informacje zwrotne od klientów służące budowie modeli danych.

W literaturze przedmiotu wskazuje się także szereg wyzwań dotyczących wykorzystania analitiky danych i big data w audycie. Dotyczą one kwestii szkolenia i wiedzy biegłych rewidentów, dostępności danych, ich istotności, integralności oraz oczekiwań organów regulacyjnych i użytkowników sprawozdań finansowych. Duże obawy związane są z kwestią bezpieczeństwa danych czy możliwości ich niszczenia (np. kontroler może utracić dane podczas ich filtrowania lub cyberataku).

Kolejnym wyzwaniem jest potencjalna zmiana struktury rynku pracy w audycie, bowiem skoro wszystko, w tym rejestrowanie, kontrolowanie i sprawdzanie danych będzie automatyczne, to do prowadzenia działalności audytowej i kontrolnej niezbędne będą mniejsze zasoby kadrowe. W obszarze zasobów ludzkich kluczowa wydaje się także konieczność postawienia na wyspecjalizowane umiejętności audytorów i kontrolerów. Branża audytorska wydaje się pozostawać w tyle, jeśli chodzi o zastosowanie analitiky danych oraz rozwój kompetencji i umiejętności pracowników w tej dziedzinie oraz podążanie za najnowszymi osiągnięciami technologicznymi<sup>33</sup>.

<sup>30</sup> M. Alles, G.L. Gray: *Incorporating big data in audits: Identifying inhibitors and a research agenda to address those inhibitors*, "International Journal of Accounting Information Systems" Volume 22/2016.

<sup>31</sup> Y.A. Hezam, L. Anthonyssamy, S.D.K. Suppiah: *Big Data Analytics and Auditing: A Review and Synthesis of Literature*, "Emerging Science Journal" nr 7/2023.

<sup>32</sup> M. Boersma: *(Data) science in financial audits*, "Compact Transparency of Information" nr 4/2018.

<sup>33</sup> A. Oussous, F.Z. Benjelloun, A. Ait Lahcen, S. Belfkih: *Big Data technologies: A survey*, "Journal of King Saud University – Computer and Information Sciences" nr 4/2018.

Biorąc pod uwagę wcześniejsze rozważania wskazać można, że technologie big data mogą zostać wykorzystane w działalności NOK, zarówno w kontroli finansowej i budżetowej, jak i w wypadku kontroli wykonania zadań lub kontroli zgodności. Zasadniczym ograniczeniem, z wyjątkiem kwestii zasobów ludzkich, jest *de facto* możliwość pozyskania informacji w formie elektronicznej od instytucji administracji publicznej. Sięganie po big data nie będzie zatem możliwe w takich obszarach, w których nie ma danych informatycznych (baz danych), obrazujących działania administracji już wytworzonych w toku normalnego funkcjonowania, a także możliwości uzyskania dostępu do nich.

Zakres informacji niezbędny najwyższym organom kontroli do rzetelnego przeprowadzenia badań zależy jest od rodzaju kontroli. W wypadku kontroli zgodności badanie polega na porównaniu stanu faktycznego w jednostce kontrolowanej z wymogami opisanymi w przepisach prawa powszechnie obowiązującego, zasadach, procedurach i politykach kontrolowanego podmiotu, zasadach dobrej praktyki lub wymaganiach dotyczących samego systemu IT służącego do obsługi danych (mechanizmy kontrolne systemu). Dotyczy ono takich cech jak: kompletność, poprawność formalna, terminowość, klasyfikacja, powiązanie, jednoznaczność, ciągłość, niesprzeczność, ślad kontrolny.

Kontrola wykonania zadań co do zasady nastawiona jest na pomiar wkładu, produktów, rezultatów i oddziaływań oraz na ocenę oszczędności, wydajności i skuteczności działań. Oznacza to wykorzystanie takich analiz ekonometrycznych i statystycznych, jak: analizy wskaźnikowe; analiza koszt–korzyść, analiza opłacalności koszt–efekt); analiza korelacji; analiza regresji; analiza szeregów czasowych; analiza porównawcza (ang. *benchmarking*); analiza obwiedni danych (ang. *Data Envelopment Analysis*)<sup>34</sup>. Możliwe jest także badanie, czy systemy informacyjne generują stosowne, wiarygodne i aktualne informacje pozwalające na monitorowanie stopnia realizacji zadań.

## Podsumowanie

Uczenie maszynowe, big data czy ostatnio sztuczna inteligencja stają się czynnikami wymuszającymi zmiany w działalności wszystkich instytucji wykorzystujących informacje i dane. Sięganie po te narzędzia w sektorze publicznym, zwłaszcza przez najwyższe organy kontroli wymaga jednak odpowiedniego zaplanowania i wdrożenia metodyki. Największym wyzwaniem wydaje się tu jednak kwestia zapewniania odpowiednio przeszkolonej kadry.

Możliwe jest zastosowanie dwóch modeli wykorzystania narzędzi big data w kontroli NOK: stworzenie wyspecjalizowanego zespołu wspierającego kontrolerów w analityce danych bądź uznanie,

---

<sup>34</sup> W. Karliński: *Metody i narzędzia wspomagania informatycznego kontroli stosowane w NIK [w:] Zastosowanie narzędzi wspomagania informatycznego kontroli (CAATs) w działalności kontrolnej i audytorskiej instytucji sektora publicznego w Polsce. Materiały z Konferencji „Przegląd Metodyczny – numer specjalny”*, Najwyższa Izba Kontroli, kwiecień 2016.

że takie kompetencje powinien mieć kontroler. Niezbędne jest także zbudowanie dużego zasobu wiedzy (informacji i danych) na potrzeby analiz służących kontrolom. Źródłem dla tego zasobu mogą być zarówno informacje pozyskane od podmiotów kontrolowanych, jak i powstałe w wyniku łączenia danych z różnych baz, co pozwala uzyskać wiedzę większą niż wynikającą z sumy informacji podlegających łączeniu.

Stosowane w sektorze publicznym narzędzia big data pozwalają także skutecznie zapobiegać wystąpieniu ryzyka (nieprawidłowości – w wypadku domeny najwyższych organów kontroli), co jest możliwe dzięki stosowaniu modeli predykcyjnych, pozwalających wskazać na wysokie prawdopodobieństwo wystąpienia nieprawidłowości zanim ona jeszcze nastąpi.

Narzędzia big data znajdują zastosowanie w obszarze finansów (publicznych) czy zamówień publicznych. Ich potencjał może zostać wykorzystany także w kontroli wykonania zadań, bowiem właśnie tu możliwe jest wyjście poza analizę transakcji i obiegu dokumentów oraz wdrożenie analityki danych pozwalającej na ocenę działalności homogenicznych podmiotów.

dr MARCIN BĘDZIESZAK,  
główny specjalista k.p.,  
Wydział Analiz Strategicznych  
Departament Strategii NIK,  
Szkoła Główna Handlowa w Warszawie  
Instytut Ekonomii Politycznej,  
Prawa i Polityki Gospodarczej  
ORCID: 0000-0003-3167-6356

**Słowa kluczowe:** analityka danych, big data, duże zbiory danych, cyfrowa transformacja w NOK

### Bibliografia:

1. Agrawal A., Gans J., Goldfarb A.: *Prediction machines: the simple economics of artificial intelligence*, Harvard Business Press Boston, MA 2018.
2. Alles M., Gray G.L., *Incorporating big data in audits: Identifying inhibitors and a research agenda to address those inhibitors*, International Journal of Accounting Information Systems, Volume 22/2016.
3. Bartoszewicz M.: *Chiński system zaufania społecznego*, „Przegląd Geopolityczny” nr 32/ 2020.
4. *Big data: The next frontier for innovation, competition, and productivity*, McKinsey Global Institute, 2011.
5. Boersma M.: *(Data) science in financial audits*, Compact Transparency of Information 2018/4.
6. Bozhinovska Lazarevska Z., Tocev T., Dionisijev I.: *How to Improve Performance in Public Sector Auditing Through the Power of Big Data and Data Analytics? – The Case of The Republic of North Macedonia*, “Journal of Accounting, Finance and Auditing” Issue 3/2022.
7. Cox M., Ellsworth D.: *Managing Big Data for Scientific Visualization*, Los Angeles 1997.

8. *European Data Market Study 2021 2023, D2.4 Second report on facts and figures*, IDC-Lison Research Council, European Commission, luty 2023.
9. Hezam Y.A., Anthonysamy L., Suppiah S.D.K., *Big Data Analytics and Auditing: A Review and Synthesis of Literature*, "Emerging Science Journal", nr 7/2023.
10. Karliński W.: *Metody i narzędzia wspomagania informatycznego kontroli stosowane w NIK [w:] Zastosowanie narzędzi wspomagania informatycznego kontroli (CAATs) w działalności kontrolnej i audytorskiej instytucji sektora publicznego w Polsce. Materiały z Konferencji, „Przegląd Metodyczny – numer specjalny”, Najwyższa Izba Kontroli, Warszawa 2016.*
11. Kelly R.: *How data analytics can help with audits.*
12. Kogosowski M.: *Israel State Comptroller to Audit Country's Readiness for AI Era.*
13. Kosior K.: *Big data w sektorze publicznym – szanse, ograniczenia, perspektywy*, „Kultura i Polityka” nr 20/2016.
14. Maciejewski M.: *Skuteczność i efektywność kontroli administracji - zbiory danych (big data) i możliwość ich wykorzystania*, „Kontrola Państwowa” nr 5/2015.
15. Munne R.: *Big Data in the Public Sector [w:] New Horizons for a Data-nDriven Economy A Roadmap for Usage and Exploitation of Big Data in Europe*, J.M. Cavanillas, E. Curry, W. Wahlster (red.), Springer 2016.
16. Otia J. E., Bracci E.: *Digital transformation and the public sector auditing: The SAI's perspective*, "Financial Accountability & Management" nr 2/2022.
17. Oussous A., Benjelloun F.Z., Ait Lahcen A., Belfkih S.: *Big Data technologies: A survey*, "Journal of King Saud University – Computer and Information Sciences" nr 4/2018.
18. Pańkowska M., Żytniewski M.: *Modele architektury biznesowej administracji publicznej w warunkach przetwarzania danych masowych*, „Roczniki Kolegium Analiz Ekonomicznych” nr 56/2019.
19. Sharma H.: *Look for ways to adopt AI into audit techniques for effectiveness: CAG.*
20. *The Path to Becoming a Data-Driven Public Sector*, OECD Digital Government Studies, OECD Publishing 2019.
21. Tomar L., Guicheney W., Kyarisiima H., Zimani T., Roseth B., S. Acevedo: *Big Data in the Public Sector. Selected Applications and Lessons Learned*, Institutions for Development Sector – Institutional Capacity of the State Division, Discussion Paper N°Idb-Dp-483, Inter-American Development Bank, October 2016.
22. *Use of big data and AI in fighting corruption and misuse of public funds. Good practice, ways forward and how to integrate new technology into contemporary control framework, organized by the Policy Department on Budgetary Affairs for the Committee on Budgetary Control*, European Parliament 2021.
23. Veenstra van A.F., Grommé F., Djafari S.: *The use of public sector data analytics in the Netherlands*, "Transforming Government: People, Process and Policy" nr 4/2021.
24. W.R. Titera: *Updating Audit Standard--Enabling Audit Data Analysis*, "Journal of Information Systems" nr 1/2013.
25. Wieczorkowski J.: *Wykorzystanie koncepcji big data w administracji publicznej*, „Roczniki Kolegium Analiz Ekonomicznych” nr 33/2014.

26. Woolley S.C., Howard P.N.: *Automation, Algorithms, and Politics | Political Communication, Computational Propaganda, and Autonomous Agents — Introduction*, "International Journal of Communication" nr 0 (October 12, 2016).

#### ABSTRACT

##### **Big Data and Digital Transformation of SAIs – Institutional Framework, Benefits and Challenges**

Over the last years, we have been witnessing a fast development of data processing, starting from machine learning, through big data, to artificial intelligence (AI), which results in new opportunities for using them in many areas (e.g. security, finance, trade, marketing). The European Commission estimates that the share of data based economy in the EU gross domestic product reached 3.9 percent in 2022. The area where big data has been gradually used is the activity of Supreme Audit Institutions (SAIs). There are, however, many challenges here, and the need to introduce methodological and organisational changes. Yet this may lead to more effective auditing of the public administration and other entities, and to better spending of resources. The aim of the article is to discuss opportunities of using big data tools in the activities of Supreme Audit Institutions. The first part of the article presents the significance of this technology in the public sector, the second part discusses the conditions related to using big data, and the last part contains considerations related to digital transformation of SAIs, as well as factors that have an impact on the use of big data in SAIs.

**Marcin Będzieszak, PhD**, senior public audit expert, Strategic Analyses Unit, Department of Strategy, NIK; Warsaw School of Economics, Institute of Political Economy, Law and Economic Policy, ORCID: 0000-0003-3167-6356

**Key words:** data analytics, big data, digital transformation in SAIs