Małgorzata Pańkowska[1], Mariusz Żytniewski[2]

Modele architektury biznesowej administracji publicznej w warunkach przetwarzania danych masowych

Streszczenie
Urzędy administracji publicznej stanęły wobec wyzwania jakim jest *smart city*, czyli inteligentne miasto lub inteligentna gmina. Celem artykułu jest prezentacja modelu architektury inteligentnego miasta, ze wskazaniem rozwiązań w technologii *Big Data*. Został również przedstawiony przykład odkrywania wiedzy na temat procesów biznesowych na podstawie logów rejestrowanych w nieustrukturalizowanych bazach danych.
Słowa kluczowe: dane masowe, internet rzeczy, inteligentne miasto, administracja publiczna, Miasto 4.0, eksploracja procesu

1. Wprowadzenie

Implementacja technologii informacyjno-komunikacyjnych w gminie jest uwarunkowana dostępnością infrastruktury szerokopasmowego internetu i usług elektronicznych. W przypadku spełnienia tych warunków, kwestią zasadniczą pozostaje ustalenie, czy wykorzystanie informacji przetwarzanych w ramach infrastruktury inteligentnych miast przez podmioty administracji publicznej i powiązane z nimi jednostki biznesowe podlega przepisom na gruncie prawa krajowego i europejskiego. Jeśli i te kwestie są rozstrzygnięte, wdrożone technologie informacyjno-komunikacyjne umożliwiają gromadzenie i przetwarzanie danych z różnych źródeł i wprowadzanie rozwiązania technicznego na rynek i do publicznego stosowania. W artykule zostanie przedstawiony model architektury administracji publicznej powodowany implementacją technologii *Big*

[1] Uniwersytet Ekonomiczny w Katowicach, Wydział Informatyki i Komunikacji, malgorzata.pankowska@ue.katowice.pl.
[2] Uniwersytet Ekonomiczny w Katowicach, Wydział Informatyki i Komunikacji, mariusz.zytniewski@ue.katowice.pl.

Data. W dalszej części artykułu znajdzie się prezentacja elementów teorii dotyczącej odkrywania wiedzy z danych przechowywanych w systemach *Big Data*. Ukazany zostanie przykład zastosowania wybranych algorytmów wspomagających odkrywanie przebiegu procesu biznesowego na bazie danych generowanych przez systemy obsługi mieszkańca. W części końcowej są przedstawione wnioski dotyczące transformacji jednostek administracji publicznej ze względu na rozwój technologii *Big Data*.

2. Rozwój inteligentnego miasta w kierunku Miasta 4.0

Studia literatury przedmiotu prowadzą do spostrzeżeń, że rozwój inteligentnego miasta (*smart city*) postępuje w obszarach, takich jak: zarządzanie siecią energetyczną w czasie rzeczywistym dla oszczędności i likwidacji strat energii i dla dostarczania jej w efektywny sposób, zarządzanie systemami transportowymi dla poprawy drożności komunikacji miejskiej, monitorowanie sieci wodociągowych i ochrony środowiska, a także zarządzanie usługami dla mieszkańców, np. e-urząd, e-zdrowie, e-edukacja i e-turystyka[3]. Internet rzeczy, przetwarzanie w chmurze, przetwarzanie danych masowych wymuszają zmianę architektury systemów administracji publicznej. Wiele urządzeń komputerowych, czujników, sensorów połączonych w sieci oraz duże ilości danych są gromadzone, przechowywane i analizowane w celu kontroli i podejmowania decyzji. Jednostka administracji publicznej zmienia swój model biznesowy i staje się organizacją komunikacji omnikanałowej[4], w której partnerzy biznesowi rozwijają swoje relacje i wymianę informacji w wielu kanałach, niejednokrotnie równolegle i zachowując przy tym swobodę wyboru kanału komunikacyjnego. Omnikanałowość przyczynia się do pewnej redundancji komunikatów, ale jednocześnie intensyfikuje działania zintegrowanych łańcuchów wartości. O ile koncepcja inteligentnego miasta była zorientowana na rozwój internetu rzeczy i oprogramowania urządzeń mobilnych, o tyle idea Miasta 4.0 rozwija się podobnie jak koncepcja czwartej rewolucji przemysłowej (*Industrie 4.0*), czyli bazując na wzajemnej zależności i zintegrowanym wykorzystaniu automatyzacji, przetwarzania i wymiany

[3] A. Monarcha-Matlak, *The Cloud Over the Smart City*, w: *Internet rzeczy, Bezpieczeństwo w Smart City*, red. G. Szpor, Wydawnictwo C.H. Beck, Warszawa 2015, s. 173–182.

[4] P. Weill, S.L. Woerner, *What's Your Digital Business Model? Six Questions to Help You Build the Next-generation Enterprise*, Harvard Business Review Press, Boston, 2018.

danych dla administracji publicznej[5]. Internet rzeczy (IoT), bazy danych, duże ilości danych, przetwarzanie w chmurze wyznaczają projekty informatyczne dla administracji publicznej.

Jednostki administracji publicznej starają się dostrzegać korzyści technologii, która pozwala na integrację agencji rządowych i wykorzystania *Big Data*, internetu rzeczy, usług dla mieszkańców oraz cyfrowej transformacji i zmian zasobów ludzkich, a także procesów pracy[6]. Rewolucja przetwarzania danych dotycząca ilości danych dostępnych i techniczne innowacje mogą pomóc w walce z korupcją. Technologia *Big Data* ma być stosowana dla tworzenia transparentności w organizacji biznesowej przez zwiększenie automatyzacji i dokładności procesów. Technologia ta może być zaimplementowana w obszarach ochrony zdrowia publicznego, handlu, podatków, w których predyktywna analiza i wizualizacja determinują trendy, wzorce i relacje w danych masowych, gdzie możliwy jest kontroling procesów administracji publicznej i monitorowanie postępowań przetargowych[7]. Wbrew wyłaniającym się słabościom, decydenci polityczni, administracja publiczna i mieszkańcy poszukują sposobów, by technologia *Big Data* stała się użyteczna dla doskonalenia efektów działania sektora publicznego, polityki publicznej, programów i procesów. W aspekcie zastosowań, najlepiej byłoby charakteryzować dane masowe (*Big Data*) jako proces, który jest inicjowany ze względu na powszechną dostępność i działanie urządzeń rejestrujących, czyli komputerów i urządzeń monitorujących, nawet jeśli są one rozproszone w niezależnych lokalizacjach. Jednostki administracji publicznej przeznaczają swój czas pracy na generowanie, gromadzenie, przetwarzanie i magazynowanie informacji i danych, dlatego że stanowią one podstawę podejmowania przez te jednostki decyzji, co określa istotę ich działania.

Pojawienie się danych masowych (*Big Data*) powoduje zmianę strategii zarządzania danymi, czyli integracji, doskonalenia jakości danych oraz gospodarowania nimi dla efektów ekonomicznych. Zanim dane zostaną użyte w analityce biznesu konieczne jest właściwe uwzględnienie kwestii, takich jak wzajemna zgodność i niesprzeczność danych, jakość, bezpieczeństwo, prywatność, przygotowanie danych i analiza. Integracja danych wymaga identyfikacji źródeł danych, jak i samych

---

[5] *Government 4.0: State-ups in Europe*, www.cebit.de/en/news-trends/news/government-4–0-state-ups-in-europe-1462.

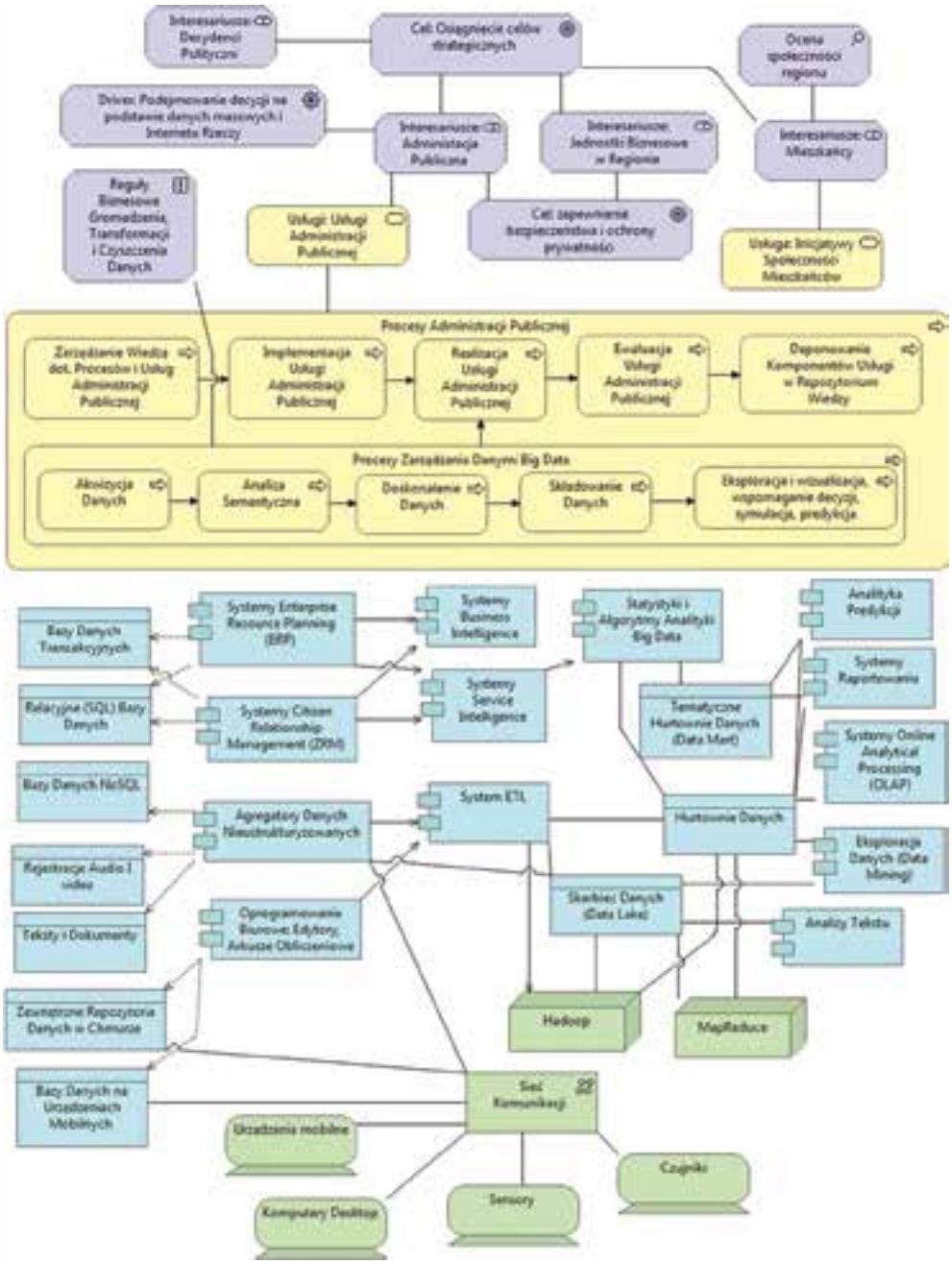
[6] D. Sagarik, P. Chansukree, W. Cho, E. Berman, *E-government 4.0 in Thailand: The Role of Central Agencies*, „Information Polity" 2018, 23, s. 343–353, content.iospress.com/download/ information-polity/ ip180006?id=information-polity%2Fip180006.

[7] L. Silveira, *Technologies Helping Us to Fight Corruption*, www.weforum.org/agenda/2016/04/4-technologies-helping-us-to-fight-corruption/.

danych, definiowania kluczy dla łączenia danych z różnych źródeł, rozpoznania luk w danych, oceny dostępności danych w czasie i dla odpowiednich użytkowników. Zarządzanie jakością danych polega na sprawdzaniu atrybutów danych, takich jak dokładność danych i uwolnienie od błędów, kompletność i pochodzenie danych z wymaganych źródeł, zgodność i aktualność danych, czytelność, elastyczność agregacji i dostępność dla przetwarzania zarówno w procesie ETL (*Extract Transform Load*), jak i w procesie ELT (*Extract Load Transform*). Bezpieczeństwo i prywatność danych są osadzone w kontekście organizacji biznesowej, który może być rozumiany jako kontekst czasu, miejsca, użycia i kultury danej organizacji społecznej. Dla jednostek administracji publicznej dane masowe są strategicznym zasobem, dlatego gospodarowanie nimi (*Data Governance*) wymaga metod analityki (*Data Science*) oraz odpowiedniej architektury *Big Data*.

3. Model architektury Miasta 4.0

Przez pojęcie Miasto 4.0 należy rozumieć jednostkę administracji publicznej, która wykorzystuje do przetwarzania danych masowych zarówno serwery w chmurze, jak i technologie mobilne internetu rzeczy oraz roboty zastępujące człowieka w realizacji usług społecznych. Dla takiej organizacji administracji publicznej zachodzi konieczność zaprojektowania architektury systemowej, obejmującej architekturę biznesową, czyli projekty struktur organizacyjnych i procesów biznesowych, architekturę danych i oprogramowania oraz architekturę techniczną, czyli sprzęt komputerowy i urządzenia rejestracji danych. Projektowanie architektury danych wymaga ustalenia systemów źródłowych i sensorów, mechanizmów transferu i procesów czyszczenia danych, końcowych baz danych i aplikacji wizualizacji danych. Wymienione decyzje są krytyczne i jako takie pozostają zależne od strategii tej organizacji społecznej oraz rozwiązań dostępnych na rynku teleinformatycznym. Na rysunku 1 przedstawiono model architektury dla przetwarzania danych masowych. Jego poszczególne warstwy odnoszą się do problemów strategicznych, technologii danych, oprogramowania i systemów komputerowych. Zasadniczym celem rozwoju architektury *Big Data* jest implementacja takiego modelu architektury, który zapewni stosowanie metod i narzędzi analityki danych (*Data Science*) i pozwoli na gospodarowanie danymi masowymi w cyklu obejmującym obserwację, tworzenie, rejestrację, zachowanie na nośnikach pamięci, zabezpieczenie, przeglądanie, współdzielenie, użycie do analizy oraz likwidację w sytuacji braku zasadności dalszego użycia.



Rysunek 1. Model architektury biznesowej i technicznej dla przetwarzania danych masowych

Źródło: opracowanie własne.

4. Koncepcja *Big Data* i jej znaczenie w przetwarzaniu wiedzy

Przetwarzanie danych masowych (*data-intensive science paradigm*) stawało się w ostatnich latach kluczowym paradygmatem nauki, obok paradygmatu empirycznego, teoretycznego i obliczeniowego[8]. Jednym z obszarów rozwoju analityki danych masowych jest odkrywanie wiedzy w danych[9]. W zakresie teoretycznym odkrywanie wiedzy dotyczy metod i algorytmów wspomagających np. odkrywanie zależności, reguł w danych lub budowanie modeli predykcyjnych. Efektem ich zastosowania jest pozyskiwanie wiedzy o świecie i zjawiskach, jakie w nim następują. Zastosowanie metod odkrywania wiedzy w danych nadaje wartość przetwarzanym w organizacji danym[10].

Podstawą koncepcji *Big Data* są dane, które w zakresie administracji publicznej mogą być rozpatrywane w dwóch ujęciach. Z jednej strony mogą to być dane dotyczące osób korzystających z usług oferowanych przez administrację publiczną. Wówczas takie dane mogą posłużyć do lepszego rozumienia potrzeb mieszkańców, określenia jakości oferowanych usług, wskazania zmian w prawie oraz procesach realizowanych przez jednostki administracji publicznej. Z drugiej strony dane mogą być generowane przez same jednostki administracji publicznej. Wówczas ich zastosowanie będzie dotyczyło: poprawy procesów, jakie zachodzą w jednostkach administracji publicznej, budowy mechanizmów bezpieczeństwa opartych na profilowaniu działań pracowników, ułatwienia dostępu do danych uczestników procesów, poprawy transparentności działań podejmowanych przez jednostki samorządu terytorialnego.

Big Data powinno być rozpatrywane nie tylko przez pryzmat dużej ilości danych, ich różnorodności oraz tempa ich generowania (traktowane jako 3V – *Volume*, *Variety*, *Velocity*), ale także w zakresie rachunku ekonomicznego wynikającego z jego zastosowania (*Value*)[11] oraz jakości i zaufania do procesów, jakie w nim zachodzą (*Veracity*)[12]. W takim ujęciu kluczowe stają się nie tylko

---

[8] A.J.G. Hey, S. Tansley, K. Tolle, *The Fourth Paradigm: Data-intensive Scientific Discovery*, Microsoft Research, 2009.

[9] S. Scoville, F. Mubarak, M. Tuohimaa, M. Zhan, *Big Data in the Public Sector: A Systematic Literature Review*, „Scandinavian Journal of Public Administration" 2017, 21(3).

[10] Ibidem.

[11] C.L. Philip Chen, C.-Y. Zhang, *Data-intensive Applications, Challenges, Techniques and Technologies: A Survey on Big Data*, „Information Sciences" 2014, 275, s. 314–347.

[12] S. Fosso Wamba, S. Akter, A. Edwards, G. Chopin, D. Gnanzou, *How 'Big Data' Can Make Big Impact: Findings from a Systematic Review and a Longitudinal Case Study*, „International Journal of Production Economics" 2015, 165, s. 234–246.

technologie informatyczne wspierające przetwarzanie danych, ale także rozwiązania analityczne wspomagające ich analizę i wnioskowanie.

W przypadku *Big Data* najczęściej stosowanymi są analizy statystyczne, metody optymalizacji oraz metody odkrywania wiedzy. Charakterystyczne dla teorii *Big Data* jest ukierunkowanie na ciągły przyrost danych, które podlegają rejestracji, a następnie analizie. Powoduje to, że klasyczne metody klasyfikacji, regresji, grupowania, asocjacji – znane z teorii *Business Intelligence* – mogą być niewystarczające do analizy ciągle zmieniających się danych. W szczególności danych na temat procesów biznesowych, którymi mogą być ciągi zdarzeń rejestrowanych przez różne systemy informatyczne. W takim ujęciu konieczne staje się zastosowanie nowych teorii z obszaru odkrywania wiedzy, jakim jest odkrywanie wiedzy na temat procesów (*proces mining*)[13].

5. Przykład odkrywania wiedzy na temat procesów obsługi mieszkańca

Modelowanie procesów biznesowych polega na ciągłej analizie procesów zachodzących w badanym podmiocie oraz określaniu wzorców możliwych czynności podejmowanych przez podmioty zaangażowane w proces. Odkrywanie wiedzy na temat procesów jest pojęciem szerszym niż modelowanie procesów, ponieważ zakłada ono eksplorację danych dostępnych w systemach informatycznych, urządzeniach, internecie w celu modelowania procesów biznesowych, monitorowania ich przebiegu oraz doskonalenia ich działania[14].

Odkrywanie wiedzy na temat procesów jest ukierunkowanie na trzy podstawowe działania: odkrywanie procesów, weryfikację modelu procesu oraz udoskonalanie istniejących procesów. Dzięki temu możliwe staje się monitorowanie i doskonalenie zachowania organizacji lub pojedynczych procesów. Jednak nie każde procesy są możliwe do łatwego zdiagnozowania oraz zamodelowania. Przykładem są tutaj analizy sieci społecznościowych oraz oddziaływanie na siebie ich uczestników. Innym przykładem zastosowania analizy procesów są analizy odchyleń od wzorcowych procesów, jakie powinny być realizowane w podmiotach, w których te procesy zachodzą. W takim ujęciu metody

[13] W.M.P. van der Aalst, *Process Mining: Discovery, Conformance and Enhancement of Business Processes*, Springer, 2011.

[14] W.M.P. van der Aalst, A. Adriansyah, A.K.A. de Medeiros i in., *Process Mining Manifesto*, Springer, Berlin, Heidelberg 2012, s. 169–194.

odkrywania procesów mogą pozwolić na poprawę jakości realizowanych procesów oraz wspomóc wykrycie anomalii wpływających na działanie organizacji. Innym obszarem jest tutaj wsparcie automatycznego budowania procesów biznesowych oraz w dalszym etapie symulacja ich przebiegu.

W literaturze odkrywanie wiedzy na temat procesów odnosi się do czerech głównych aspektów[15]:

- **analizy przebiegu procesu** – gdzie są analizowane czynności oraz kolejność ich realizacji w celu określenia modelu procesu wynikającego z danych faktycznie zebranych,
- **analizy uczestników i zasobów** – ma to na celu określenie zasobów ludzkich i rzeczowych zaangażowanych w proces,
- **analizy instancji** – dotyczy ona danych związanych z faktycznym przebiegiem pojedynczych instancji procesu,
- **analizy dopasowania modelu procesu do jego faktycznego przebiegu**.

Takie odniesienie do analizy procesów biznesowych pomija jednak zasadniczy aspekt jakości wyników analiz, jakie mogą być generowane przez różne algorytmy. Jest to szczególnie ważne w zakresie odkrywania przebiegu procesu biznesowego, ponieważ działanie to zakłada, iż projektowanie procesu będzie oparte jedynie na danych zebranych w logach procesów. W takim ujęciu konieczne jest określenie parametrów oceny odkrytego procesu biznesowego. Aalst[16] wskazuje na cztery parametry:

1. **Dopasowanie (*fitness*)** – określa, jak model procesu odpowiada danym zawartym w logu procesu oraz jak kolejne instancje procesu będą do tego modelu dopasowane.
2. **Prostota (*simplicity*)** – określa prostotę procesu zgodnie z koncepcją brzytwy Ockhama. Odkryty model procesu powinien być możliwie jak najprostszy, bez utraty wiedzy na temat zdarzeń, które wchodzą w jego skład.
3. **Uogólnienie (*generalization*)** – miara ta wskazuje, jak dobrze model będzie odtwarzał przyszłe zdarzenia. Przykładami modeli o niskiej wartości tego wskaźnika są modele o powtarzających się sekwencjach zdarzeń ujmowanych w osobnych ścieżkach przepływu modelu. Modele takie charakteryzują

[15] W. Fliegner, *Metoda analizy międzyorganizacyjnych procesów biznesowych wspomaganych komputerowo*, „Studia Ekonomiczne. Zeszyty Naukowe Uniwersytetu Ekonomicznego w Katowicach" 2015, 224, s. 28–40.

[16] W.M.P. van der Aalst, *Process mining*…, op. cit.

się również dużą obojętnością. W literaturze wskazuje się, że jest to miara zaufania do precyzji modelu[17].

4. **Precyzja (*precision*)** – określa wskaźnik występowania zdarzeń dozwolonych przez model, ale niezdiagnozowanych w logu. Im ten wskaźnik jest wyższy, tym mniej sekwencji zdarzeń, na które pozwala model, a które wykraczają poza dane źródłowe. Przykładem modelu o niskiej precyzji jest model kwiatu[18].

W przypadku pierwszej miary dopasowania może być ona rozpatrywana w trojaki sposób (miary te znajdują swoje zastosowanie w ocenie jakości odkrytego modelu w oprogramowaniu ProM[19]):

1. **Ogólna miara dopasowania modelu i danych (*trace fitness* – TF)** – określa dopasowanie zaproponowanego modelu do danych, z których został on wygenerowany. Wartość 0 oznacza brak dopasowania modelu do danych. Wartość 1 oznacza pełne dopasowanie danych do modelu.
2. **Miara zgodności modelu (*move-model fitness* – MMF)** – wartość poniżej 1 wskazuje, że istnieje niezgodność między zaobserwowanym zdarzeniem w logu procesu a modelem zdiagnozowanym przez dany algorytm. Zgodnie z opracowanym modelem zdarzenie było oczekiwane, jednak w logu procesu nie zostało zaobserwowane.
3. **Miara zgodności danych (*move-log fitness* – MLF)** – przeciwieństwo MMF. W tym wypadku w logu procesu pojawiło się zdarzenie, które nie zostało zaobserwowane w opracowanym modelu.

Przykład analizy dotyczy procesu realizacji sprawy mieszkańca i odnosi się do danych pochodzących z systemu informatycznego w zakresie 200 zewidencjowanych spraw oraz ponad 1000 zdarzeń.

We wskazanym przykładzie zastosowano algorytmy odkrywania wiedzy o przebiegu procesu biznesowego AlphaPlus[20] oraz AlphaSharp[21]. Wyniki analizy wskazanych danych prezentuje tabela 1.

---

[17] J.C.A.M. Buijs, B.F. van Dongen, W.M.P. van der Aalst, *On the Role of Fitness, Precision, Generalization and Simplicity in Process Discovery*, w: *On the Move to Meaningful Internet Systems: OTM 2012 (Confederated International Conferences: CoopIS, DOA-SVI, and ODBASE 2012*, red. R. Meersman, Rome, Italy, September 10–14, 2012, Proceedings, Part I, s. 305–322.

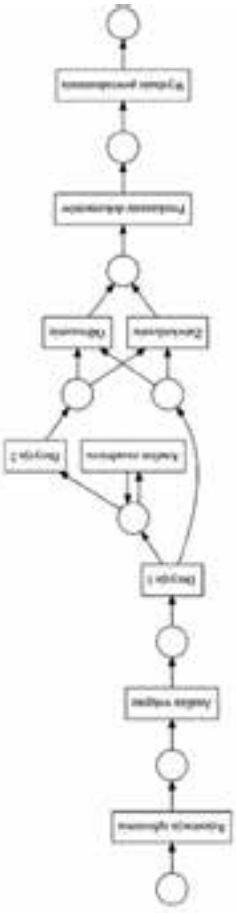
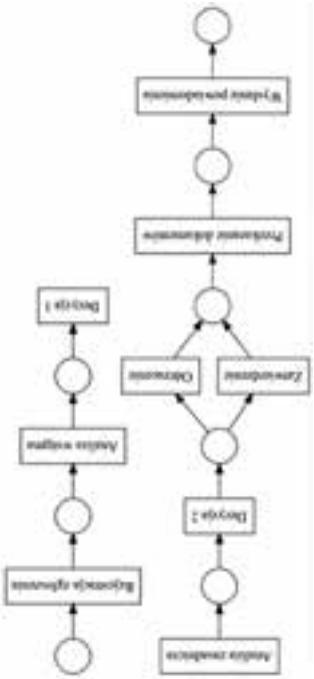
[18] A. Rozinat, W.M.P. van der Aalst, *Conformance Checking of Processes Based on Monitoring Real Behavior*, „Information Systems" 2008, 33(1), s. 64–95.

[19] www.promtools.org (data odczytu: 26.10.2018).

[20] W.M.P. van der Aalst, *Process mining*…, op. cit.

[21] L. Wen, J. Wang, W.M.P. van der Aalst, B. Huang, J. Sun, *Mining Process Models with Prime Invisible Tasks*, „Data & Knowledge Engineering" 2010, 69(10), s. 999–1021.

Tabela 1. Rezultaty zastosowania algorytmów AlphaPlus i AlphaSharp

Rodzaj algorytmu	Wygenerowany model	TF	MMF	MLF	Uogólnienie	Precyzja
AlphaPlus		**0,98**	0,96	**1,00**	0,99	**0,94**
AlphaSharp		0,57	**1,00**	0,40	0,99	0,50

TF – Miara dopasowania modelu i danych.
MFF – Miara zgodności modelu.
MLF – Miara zgodności danych.

Jak widać, wskazane miary opisu procesu biznesowego ukazują dużą przewagę algorytmu AlphaPlus w zakresie analizowanych danych. Model ten wykazał się dużą precyzją oraz dopasowaniem. Jak wskazuje wygenerowany diagram tego procesu z zastosowaniem sieci Petriego, jedno zdarzenie ma charakter rekurencyjny („Analiza zasadnicza"). Wskazuje to, że zdarzenie to może występować wielokrotnie po sobie. Może to stanowić anomalię i powinno zostać zweryfikowane z wzorcowym modelem tego procesu. W dalszym etapie tak wygenerowany model może posłużyć do symulacji jego przebiegu na podstawie uzyskanych wyników historycznych.

6. Podsumowanie i kierunki dalszych badań

Rozwój teorii oraz praktyki w obszarze koncepcji *Big Data* w zakresie jej zastosowania dla wspomagania administracji publicznej oraz mieszkańców, pozwala na polepszenie jakości oferowanych usług, poprzez integrację danych pochodzących z różnych źródeł. Jednocześnie, co nie zostało omówione w niniejszym artykule, może powodować problemy natury prawnej związane ze zbieraniem i przetwarzaniem danych umożliwiających profilowanie mieszkańców. Dlatego podejmując decyzje dotyczące rozwoju przetwarzania danych masowych, należy na nie spojrzeć z trzech punktów widzenia:

- aspekt strategiczny: Jakie praktyki i modele ramowe powinny być stosowane dla wdrażania zdecentralizowanego podejmowania decyzji przy użyciu danych masowych?
- aspekt techniczny: Gdzie i jak dane powinny być przechowywane i przetwarzane dla zapewnienia ich odpowiedniej ochrony i celowej redundancji?
- aspekt analityczny: Jakie dane będą analizowane i jakie procesy analityki zostaną zaimplementowane?

W niniejszym artykule podjęto wskazane tutaj obszary poprzez opracowanie modelu architektury biznesowej i technicznej dla przetwarzania danych masowych oraz ukazanie zastosowania wybranych algorytmów odkrywania wiedzy o procesach w celu analizy przebiegu procesu biznesowego, na bazie danych zebranych z systemu informatycznego.

Bibliografia

Buijs J.C.A.M., Dongen van B.F., Aalst van der W.M.P., *On the Role of Fitness, Precision, Generalization and Simplicity in Process Discovery*, w: *On the Move to Meaningful Internet Systems: OTM 2012 (Confederated International Conferences: CoopIS, DOA-SVI, and ODBASE 2012*, red. R. Meersman, Rome, Italy, September 10–14, 2012, Proceedings, Part I, s. 305–322.

Fliegner W., *Metoda analizy międzyorganizacyjnych procesów biznesowych wspomaganych komputerowo*, „Studia Ekonomiczne. Zeszyty Naukowe Uniwersytetu Ekonomicznego w Katowicach" 2015, 224, s. 28–40.

Fosso Wamba S., Akter S., Edwards A., Chopin G., Gnanzou D., *How 'Big Data' Can Make Big Impact: Findings from a Systematic Review and a Longitudinal Case Study*, „International Journal of Production Economics" 2015, 165, s. 234–246.

Hey A.J.G., Tansley S., Tolle K., *The Fourth Paradigm: Data-intensive Scientific Discovery*, Microsoft Research, 2009.

Monarcha-Matlak A., *The Cloud Over the Smart City*, w: *Internet rzeczy, Bezpieczeństwo w Smart City*, red. G. Szpor, Wydawnictwo C.H. Beck, Warszawa, 2015, s. 173–182.

Philip Chen C.L., Zhang Ch.-Y., *Data-intensive Applications, Challenges, Techniques and Technologies: A Survey on Big Data*, „Information Sciences" 2014, 275, s. 314–347.

Rozinat A., van der Aalst W.M.P., *Conformance Checking of Processes Based on Monitoring Real Behavior*, „Information Systems" 2008, 33(1), s. 64–95.

Scoville S., Mubarak F., Tuohimaa M., Zhan M., *Big Data in the Public Sector: A Systematic Literature Review*, „Scandinavian Journal of Public Administration" 2017, 21(3).

van der Aalst W.M.P., *Process Mining: Discovery, Conformance and Enhancement of Business Processes*, Springer, 2011.

van der Aalst W.M.P., Adriansyah W., de Medeiros A. i in., *Process Mining Manifesto*, Springer, Berlin, Heidelberg 2012, s. 169–194.

Weill P., Woerner S.L., *What's Your Digital Business Model? Six Questions to Help You Build the Next-generation Enterprise*, Harvard Business Review Press, Boston 2018.

Wen L., Wang J., van der Aalst W.M.P., Huang B., Sun J., *Mining Process Models with Prime Invisible Tasks*, „Data & Knowledge Engineering" 2010, 69(10), s. 999–1021.

Źródła internetowe

Government 4.0: State-ups in Europe, www.cebit.de/en/news-trends/news/government-4–0-state-ups-in-europe-1462 (data odczytu: 7.11.2018).

Sagarik D., Chansukree P., Cho W., Berman E., *E-government 4.0 in Thailand: The role of central agencies*, „Information Polity" 2018, 23, s. 343–353, content.iospress.com/download/information-polity/ip180006?id= information-polity%2Fip180006 (data odczytu: 7.11.2018).

Silveira L., *Technologies Helping Us to Fight Corruption*, www.weforum.org/agenda/2016/04/4-technologies-helping-us-to-fight-corruption (data odczytu: 7.11.2018).

www.promtools.org (data odczytu: 26.10.2018).

* * *

Public administration business architecture models in big data computing conditions

Abstract

Public administration agencies face the challenge of the smart city that is understood as an intelligent city or town. The paper aims to present smart city business architecture models, oriented towards Big Data technology solutions that support business goals achievement. The paper covers an example of business process knowledge discovering on logs in non-structural databases.

Keywords: Big Data, Internet of Things, smart city, Government 4.0, City 4.0, process mining