

MACIEJ KIEDROWICZ<sup>1</sup>, JAROSŁAW KOSZELA<sup>2</sup>

## Architektura systemu informatycznego wspomagającego zwalczanie przestępczości gospodarczej na przykładzie systemu IAFEC

### 1. Wstęp

Celem artykułu jest zaprezentowanie nowoczesnej architektury oraz koncepcji systemu wspomagającego analizę danych na podstawie heterogenicznych zasobów. Jest to także system wykorzystujący podejście usługowe typu SOA (ang. *Service Oriented Architecture*) dla różnorodnych metod analizy danych, bazujących na zróżnicowanych metodach i algorytmach analizy (metody analizy grafowo-sieciowej poszukiwania dróg w homogenicznych i wielowarstwowych grafach) na potrzeby wspomagania zwalczania przestępczości gospodarczej. Koncepcja i architektura systemu IAFEC (ang. *Information Analysis of Financial and Economic Crime*) powstała jako konsekwencja prac, które dotyczyły możliwości wykorzystania różnorodnych zasobów danych do przeciwdziałania przestępczości gospodarczej (głównie w obszarze finansów)<sup>3, 4, 5</sup>. Tworząc architekturę tego systemu, przyjęto założenie, że wymagane dane (w szczególności związane z rejestrami i ewidencjami publicznymi) będą dostępne bez żadnych ograniczeń formalnoprawnych. Zaproponowana architektura została przedstawiona z punktu widzenia roli „Analityka”, który odpowiada za przygotowanie danych do analizy, przeprowadzenie tej analizy oraz przedstawienie jej wyników do dalszego procedowania. Wielkość proponowanych do wykorzystania

---

<sup>1</sup> Wojskowa Akademia Techniczna w Warszawie, Wydział Cybernetyki.

<sup>2</sup> Wojskowa Akademia Techniczna w Warszawie, Wydział Cybernetyki.

<sup>3</sup> M. Kiedrowicz, *Organizacja i dostęp do heterogenicznych, publicznych zasobów danych*, w: *Projektowanie systemów informatycznych: modele i metody*, red. T. Nowicki, Z. Tarapata, WAT, Warszawa 2014.

<sup>4</sup> M. Kiedrowicz, *Rejestry i zasoby informacyjne wykorzystywane przez organy odpowiedzialne za wykrywanie i przeciwdziałanie przestępczości*, w: *Jawność i jej ograniczenia: zadania i kompetencje*, red. G. Szpor, Wydawnictwo C.H. Beck, Warszawa 2015.

<sup>5</sup> *Zarządzanie informacjami wrażliwymi. Wybrane aspekty organizacyjne, prawne i techniczne ochrony informacji niejawnych*, red. M. Kiedrowicz, WAT, Warszawa 2015.

zasobów danych oraz ich zmienność powoduje, że dotychczas wykorzystywane narzędzia i metody, służące do analizy danych, okazują się niewystarczające<sup>6</sup>. Z punktu widzenia decydentów możliwość uzyskania informacji opartej na przeanalizowaniu wszystkich potencjalnych przesłanek określonych działań daje szansę na podjęcie zdecydowanie lepszych jakościowo decyzji co do dalszego postępowania (zmniejsza się ryzyko podjęcia decyzji błędnych).

W trakcie tworzenia koncepcji systemu IAFEC jednym z podstawowych założeń była możliwość uzyskania satysfakcjonujących wyników analiz jeszcze przed zajściem przestępstwa finansowego, tak aby można było zminimalizować lub wręcz wyeliminować jego negatywne skutki.

## 2. Architektura systemu IAFEC

Architektura systemu, wspomagającego proces analizy spraw związanych z przestępstwami głównie z obszarów przestępstw finansowych, składa się z pięciu głównych podsystemów realizujących najważniejsze procesy systemu<sup>7,8</sup> (rysunek 1):

- zasoby informacyjne systemu – zajmują się głównie zgromadzeniem danych w bazach danych (relacyjnych i sieciowych<sup>9,10</sup>), pozyskiwaniem danych z zasobów rejestrów administracyjnych oraz ich importem do zasobów systemu<sup>11</sup>;

---

<sup>6</sup> A. Kaucz, M. Kiedrowicz, M. Skinder-Pik, *Gromadzenie i przetwarzanie danych mających związek ze zwalczaniem przestępczości finansowej. Zasady dostępu, ograniczenia prawne*, WAT, Warszawa 2016.

<sup>7</sup> B. Gawin, *Systemy informatyczne w zarządzaniu procesami Workflow*, Wydawnictwo Naukowe PWN, Warszawa 2015.

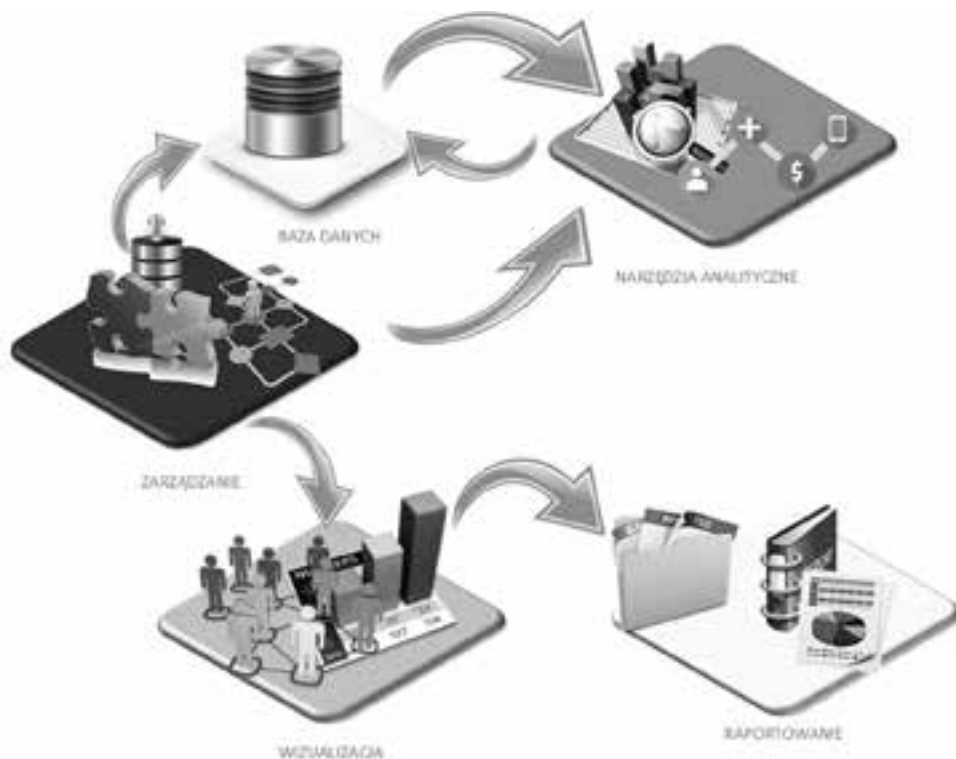
<sup>8</sup> J. Koszela, *Hurtownie procesów – systemy analitycznego przetwarzania procesów*, w: *Projektowanie systemów informatycznych: modele i metody*, red. T. Nowicki, Z. Tarapata, WAT, Warszawa 2014.

<sup>9</sup> A. Vukotic, N. Watt, T. Abedrabbo, D. Fox, J. Partne, *Neo4J in action*, Manning Publications Co., Shelter Island 2015.

<sup>10</sup> Neo4J, <https://neo4j.com/> (09.11.2016).

<sup>11</sup> M. Kiedrowicz, *Uogólniony model danych w rozproszonych rejestrach ewidencyjnych*, „Roczniki Kolegium Analiz Ekonomicznych” 2014, z. 33, s. 209–234.

- zarządzanie sprawami – pozwalające m.in. na definiowanie nowych spraw i zarządzanie istniejącymi, określanie uprawnień i dostępu, scalanie i rozdzielanie spraw<sup>12, 13,14</sup>;



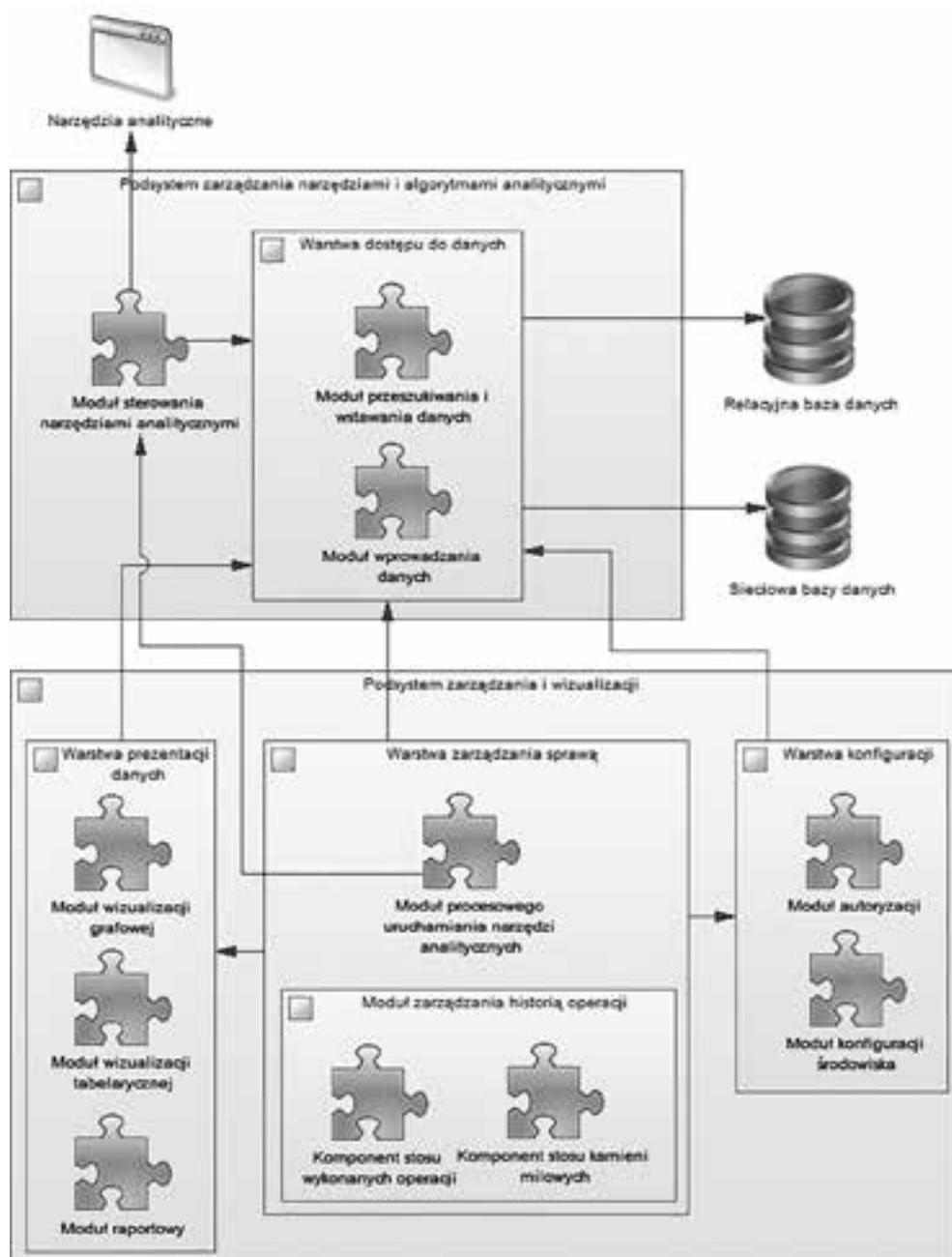
**Rysunek 1. Architektura ogólna systemu IAFEC**

Źródło: opracowanie własne.

<sup>12</sup> R. Hoffmann, M. Kiedrowicz, J. Stanik, *Risk Management System as the Basic Paradigm of the Information Security Management System in an Organization*, 20th International Conference on CSCC, MATEC Web of Conferences, vol. 76, 2016.

<sup>13</sup> M. Kiedrowicz, T. Protasowicki, J. Stanik, *Wybrane aspekty standaryzacji w ochronie publicznych zasobów informacyjnych i świadczonych usług w kontekście społeczeństwa informacyjnego*, „Zeszyty Naukowe Uniwersytetu Szczecińskiego – Ekonomiczne Problemy Usług” 2014, z. 113, s. 113–130.

<sup>14</sup> R. Hoffmann, M. Kiedrowicz, J. Stanik, *Evaluation of Information Safety as an Element of Improving the Organization’s Safety Management*, 20th International Conference on CSCC, MATEC Web of Conferences, 2016, vol. 76.



Rysunek 2. Diagram komponentów systemu wspomagającego pracę analityka

Źródło: opracowanie własne.

- narzędzia i usługi analityczne – tworzące zestaw usług analitycznych w technologii SOA (ang. *Service-Oriented Architecture*)<sup>15</sup>, wykorzystującej różnego typu metody i algorytmy analiz danych, np. algorytmy analiz semantycznych czy bazujące na teorii grafów i sieci<sup>16</sup>;
- wizualizacja spraw – pozwalająca na czytelną i spójną prezentację sprawy w postaci graficznej jako sieć powiązanych elementów oraz uzyskanych wyników usług analitycznych;
- narzędzia raportowania i eksportowania spraw – pozwalające na przygotowanie raportów z danej sprawy wraz z możliwością wyeksportowania wybranych danych o sprawie np. do plików o standardowych formatach danych dla narzędziach biurowych.

Bardziej szczegółowy model architektoniczny narzędzia wspomagającego pracę analityka w systemie IAFEC został zaprezentowany na rysunku 2. Narzędzie analityka zawiera w sobie główne funkcjonalności podsystemu zarządzania sprawami, wizualizacji, raportowania oraz komunikacji z narzędziami i usługami analitycznymi.

### 3. Zasoby informacyjne systemu IAFEC

Podsystem zasobów informacyjnych tworzą bazy danych systemu IAFEC powiązane ze sobą poprzez działające procesy ETL (ang. *Extract-Transform-Load*). Bazy te to (rysunek 3):

- zasoby i bazy rejestrowe,
- uspołniona baza zbiorcza (relacyjna),
- semantyczna baza sieciowa.

Proces gromadzenia danych rozpoczyna się od pozyskania i przetworzenia zbiorów danych wejściowych. Dane źródłowe są przetwarzane z plików (np.: csv, xml, txt) lub struktur baz danych w przypadku możliwości korzystania online z zasobów rejestrów administracyjnych. Po pozyskaniu danych źródłowych inicjowany jest proces transformacji i czyszczenia danych. Ujednoliconymi danymi jest zasilana zbiorcza relacyjna baza danych, czyli źródło danych

---

<sup>15</sup> M. Kiedrowicz, J. Stanik, *Selected Aspects of Risk Management in Respect of Security of the Document Lifecycle Management System with Multiple Levels of Sensitivity*, w: *Information Management in Practice*, eds. B.F. Kubiak, J. Maślankowski, Gdańsk 2015, pp. 231–249.

<sup>16</sup> Ł. Warchał, *Using Neo4j Graph Database in Social Network Analysis*, "Studia Informatica" 2012, vol. 33, no. 2A.

dla algorytmów i rozwiązań modułu analitycznego. Na podstawie uspojnionych danych tworzony jest graf zależności i powiązań wykorzystywany przez złożony moduł analityczny. Moduł importu danych skorzysta z następujących baz i zasobów rejestrowych:

- PESEL – dane dotyczące osób wraz ze związkami pokrewieństwa i powinowactwa,
- KRS /EDG/ REGON – dane dotyczące prowadzonych działalności gospodarczych,
- ZUS – dane płatników i osób ubezpieczonych,
- CEPIK – dane dotyczące pojazdów i ich właścicieli,
- CELNE – dane dotyczące głównie operacji celnych i gospodarczych związanych z przekraczaniem granic kraju,
- KW – własności ksiąg wieczystych podmiotów,
- BILINGI – rejestr połączeń podmiotów,
- GIIF – rejestr transakcji bankowych podmiotów.

Wykaz zasobów, z których pozyskiwane są dane do analiz zależności i powiązań, może się zmieniać w zależności od możliwości i dostępu do nowych zasobów rejestrowych krajowych i zagranicznych. Relacyjna baza danych systemu IAFEC przechowuje uspojniony, jednolity obraz danych powstały ze złączenia danych pochodzących ze wszystkich źródeł. Wraz ze zbiorczą bazą danych utrzymywana jest także baza wzorców danych oraz struktury analityczne. Moduł importu danych korzysta także z tabel parametrycznych przechowujących techniczne parametry rozwiązań.

Aby móc skutecznie zasilić powyższe bazy, jest wymagany sprawny moduł importu danych, stanowiący główną część wspomnianego podsystemu.

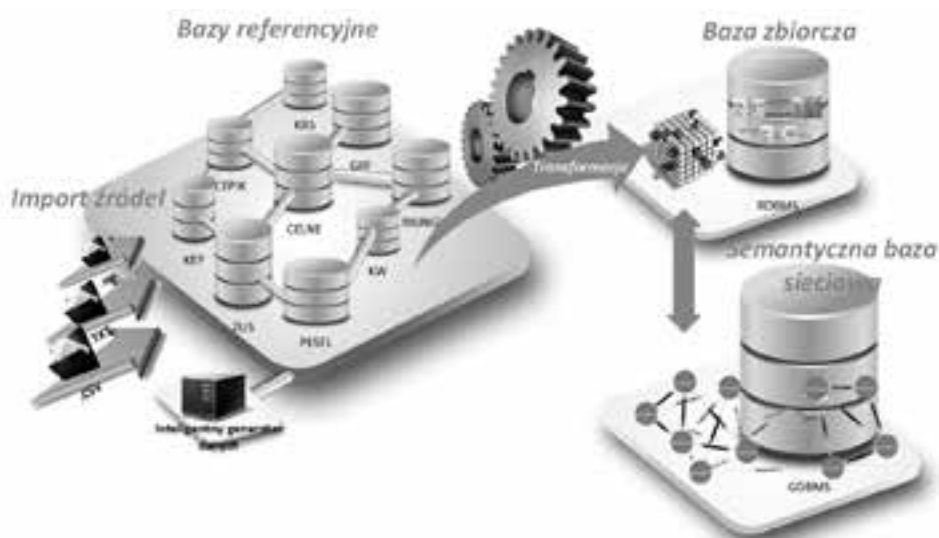
Celem modułu importu danych jest pozyskanie danych z systemów źródłowych, uspojnienie, a następnie przygotowanie odpowiednich struktur, które będą źródłem podsystemu narzędzi analiz. Podsystem zasobów informacyjnych składa się z:

- mechanizmu importującego dane z systemów źródłowych,
- kopii wycinka baz referencyjnych systemów źródłowych zawierających zasoby wykorzystywane we wnioskowaniu,
- bazy uspojnionej IAFEC zawierającej jednolite, przekształcone na potrzeby analizy dane wraz ze strukturami analitycznymi,
- graf zależności i powiązań – semantyczną bazę sieciową odwzorowującą zależności między danymi,
- inteligentny generator danych testowych.

Można wyróżnić dwa warianty dostępu do danych:

- nieograniczony – pełen dostęp do wszystkich zasobów,
- ograniczony – dostęp niepełny, wymagający przyrostowego ładowania danych.

Podstawą działania tych metod są dane gromadzone przez różne instytucje administracji w wyniku działań podmiotów i osób fizycznych związanych z przepływem środków finansowych, w tym wartości majątkowych. Dotyczy to także innych danych związanych z prowadzeniem działalności gospodarczej oraz szeroko rozumianej komunikacji pomiędzy podmiotami<sup>17</sup>.



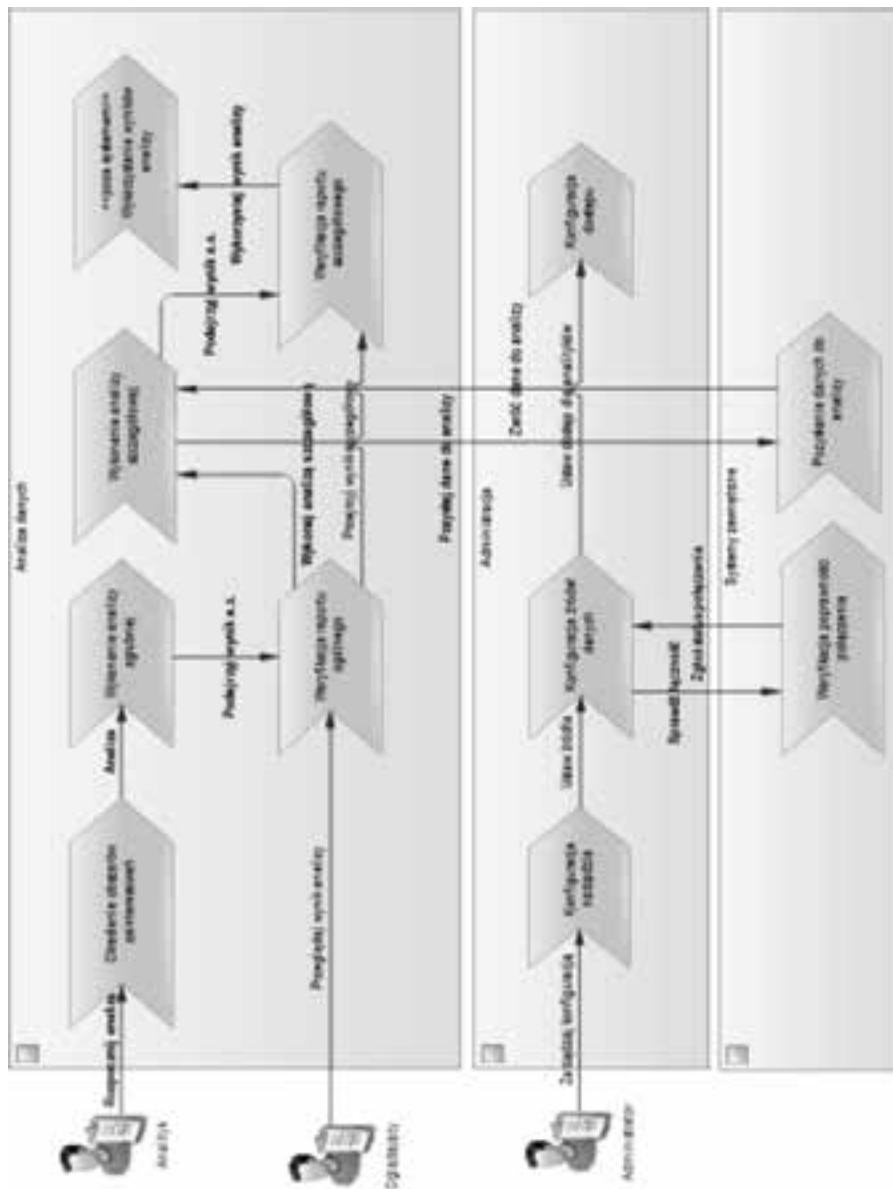
Rysunek 3. Podsystem zasobów informacyjnych

Źródło: opracowanie własne.

#### 4. Zarządzanie sprawami – proces budowy i analizy sprawy w systemie IAFEC

Za inicjowanie i zarządzanie sprawami w systemie IAFEC odpowiada podsystem zarządzania sprawami. Jego głównymi zadaniami są: zapewnienie

<sup>17</sup> M. Kiedrowicz, J. Koszela, *Business Processes Modelling for the Processing of Classified Documents Using RFID Technology*, "Collegium of Economic Analysis Annals" 2016, vol. 42, pp. 53–66.



Rysunek 4. Diagram procesu budowania sprawy przez analityka

Źródło: opracowanie własne.



intuicyjnej i efektywnej obsługi sprawy, w tym jej założenie, administrowanie oraz wspomaganie wszystkich działań wynikających z jej przebiegu. Dodatkowo podsystem musi zapewniać możliwości konfiguracji środowiska i autoryzację zgodną z powszechnie wykorzystywanymi standardami bezpieczeństwa. Kluczowe jest również umożliwienie użytkownikowi wizualizacji informacji i danych w różnych widokach oraz generowanie raportów z przebiegu sprawy. Do zrealizowania powyższych celów wydzielono trzy warstwy tego podsystemu: warstwę zarządzania sprawą, warstwę prezentacji danych oraz warstwę konfiguracji. Warstwa zarządzania sprawą zapewnia realizację funkcjonalności związanych z prowadzeniem, zakładaniem i kończeniem sprawy. W jej skład wchodzi moduł zarządzania historią operacji, który przez komponenty stosu wykonanych operacji oraz kamieni milowych pozwala na kontrolowanie procesu analizy danych i samej sprawy. Moduł procesowego uruchamiania narzędzi analitycznych przez komunikację z podsystem zarządzania narzędziami i algorytmami analitycznymi umożliwia ich wysterowanie oraz obsługę ze strony użytkownika, pozwalając mu w pełni dynamicznie kontrolować procesy analizy danych. Bazowy proces związany z budowaniem sprawy przez analityka przedstawiony jest na rysunku 4.

## 5. Narzędzie wspomagające pracę analityka

Narzędziem wspomagającym pracę analityka jest oprogramowanie, za pomocą którego można inicjować, realizować i zarządzać procesem budowy sprawy (rysunek 5)<sup>18, 19</sup>. Operowaniem na sprawie nazywamy działania o charakterze manualnym (analityk samodzielnie buduje i modyfikuje daną sprawę) oraz użycie narzędzi analitycznych w celu uzyskania dodatkowych faktów, które mogą zostać wykorzystane jako dowody albo stanowić przesłanki do dalszej i głębszej analizy<sup>20, 21</sup>.

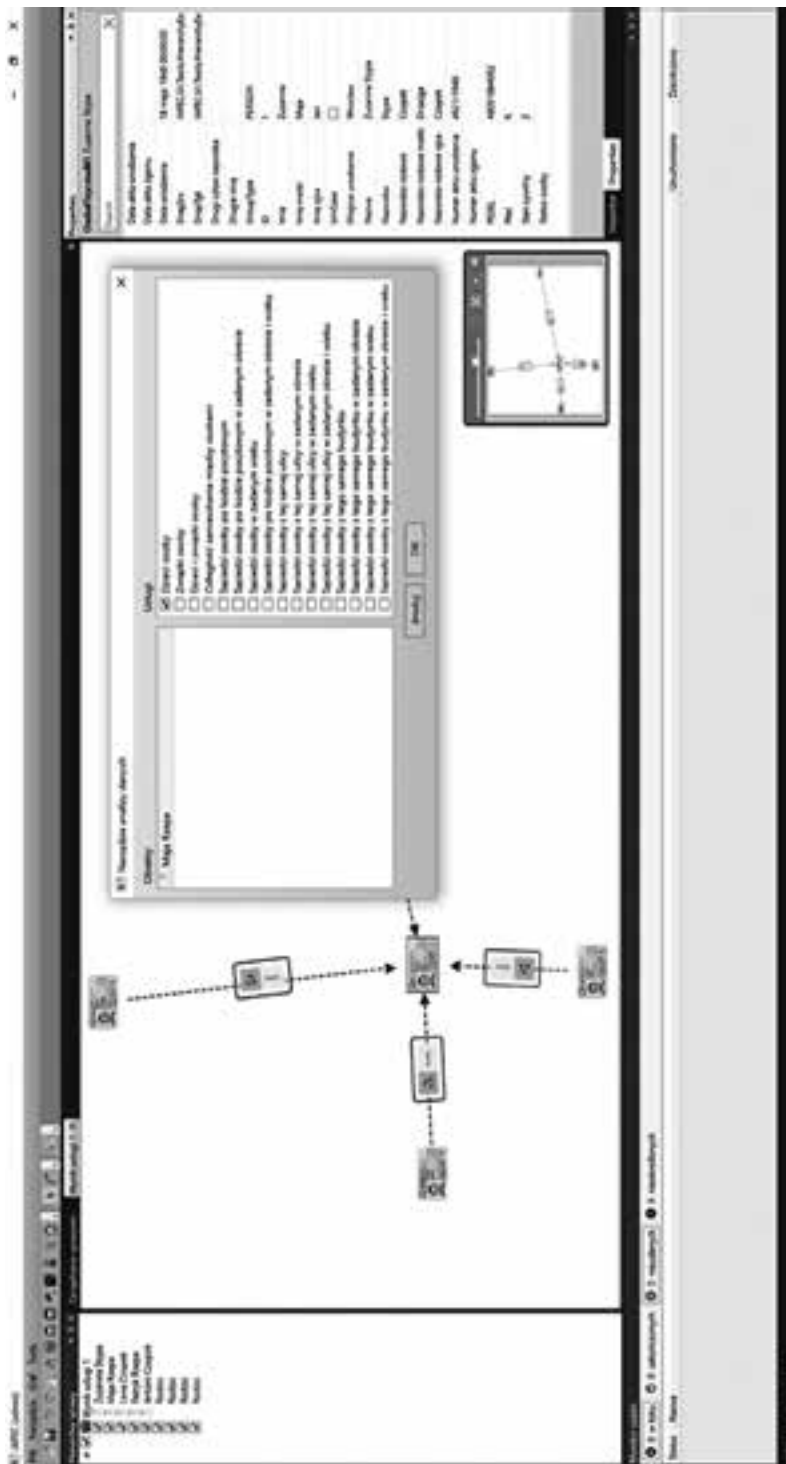
---

<sup>18</sup> M. Kiedrowicz, *Location with the Use of the RFID and GPS Technologies – Opportunities and Threats*, GIS ODYSSEY 2016, pp. 122–128.

<sup>19</sup> M. Kiedrowicz, *Objects Identification in the Informations Models Used by Information Systems*, GIS ODYSSEY 2016, pp. 129–136.

<sup>20</sup> M. Kiedrowicz, *Dostęp do publicznych zasobów danych – Big data czy Big brother*, w: INTERNET. Publiczne bazy danych i Big data, red. G. Szpor, Warszawa 2014, s. 15–39.

<sup>21</sup> M. Kiedrowicz, *Rejestry publiczne wykorzystywane przez organy odpowiedzialne za wykrywanie i przeciwdziałanie przestępczości*, w: *Rejestry publiczne: jawność i interoperacyjność*, red. A. Gryszczyńska, Warszawa 2016, s. 603–649.



Rysunek 5. Interfejs roboczy aplikacji wspomagającej pracę analityka

Źródło: opracowanie własne.

Ponadto do procesu zarządzania sprawą zaliczamy prezentację otrzymanych już przetworzonych informacji w formie zarówno grafowej (np. zobrazowanie sieci powiązań rodzinnych, społecznych), jak i tabelarycznej (np. zestawienie operacji na koncie). W ten sposób zorganizowana prezentacja danych pozwoli analitykowi, czy osobie prowadzącej sprawę, na zagłębienie się w jej tajniki w dużo większym zakresie, niż jest to możliwe przy wykorzystaniu dotychczasowych technik<sup>22</sup>. Istotnym faktem jest umożliwienie ponownego uruchomienia narzędzi i usług analitycznych na nowych zbiorach danych. Oznacza to, że dopóki sprawa nie zostanie zakończona, zbiór dotyczących jej danych zwykle narasta. Ponowne użycie tych samych metod i algorytmów analizy na zaktualizowanych danych może ujawnić nowe, istotne fakty dotyczące danej sprawy. Narzędzia analityczne są traktowane jako usługi wykorzystywane przez zestaw narzędzi IAFEC. Podsystem ten oprócz zarządzania wykorzystaniem narzędzi analitycznych i prezentacją ich rezultatów zajmuje się generowanie raportów, które mogą zostać dołączone do akt sprawy. Raporty te mogą mieć także charakter plików z wynikami poszczególnych operacji na zbiorach danych (w rozumieniu wykorzystanych metod analizy).

## 6. Podsumowanie i kierunki dalszych badań

Specyfika funkcjonowania niektórych obszarów gospodarki pociąga za sobą konieczność zastosowania odpowiednich narzędzi analitycznych, a prezentowany system i opracowane w jego ramach narzędzia mogą zostać do tego celu wykorzystane. Dalsze badania będą się skupiać na weryfikacji zaproponowanych rozwiązań i ich ewentualnej rozbudowie z wykorzystaniem pojawiających się nowych rozwiązań technologicznych. Dotyczyć to będzie również badań związanych z rozbudową modelu semantycznego zasobów danych, ze zwiększeniem wykorzystywanych wolumenów danych oraz badań i weryfikacji zastosowanych metod i algorytmów pod względem jakościowym i wydajnościowym uzyskiwanych wyników. Otrzymane wyniki posłużą również do inicjowania rozwiązań formalnoprawnych, które będą powodowały podniesienie efektywności zastosowanych narzędzi i metod, w szczególności dotyczących formalnych procedur postępowania przez służby odpowiadające za walkę z przestępczością gospodarczą.

---

<sup>22</sup> N.K. Kashyap, B.K. Pandey, H.L. Mandoria, *Analysis of Pattern Identification Using Graph Database for Fraud Detection*, Oriental Journal of Computer Science and Technology, vol. 9(2).

## Bibliografia

- Gawin B., *Systemy informatyczne w zarządzaniu procesami Workflow*, Wydawnictwo Naukowe PWN, Warszawa 2015.
- Hoffmann R., Kiedrowicz M., Stanik J., *Evaluation of Information Safety as an Element of Improving the Organization's Safety Management*, 20th International Conference on Circuits, Systems, Communications and Computers, MATEC Web of Conferences, vol. 76, Greece 2016.
- Hoffmann R., Kiedrowicz M., Stanik J., *Risk Management System as the Basic Paradigm of the Information Security Management System in an Organization*, 20th International Conference on Circuits, Systems, Communications and Computers, MATEC Web of Conferences, vol. 76, Greece 2016.
- Kashyap N.K., Pandey B.K., Mandoria H.L., *Analysis of Pattern Identification Using Graph Database for Fraud Detection*, „Oriental Journal of Computer Science and Technology”, vol. 9(2).
- Kaucz A., Kiedrowicz M., Skinder-Pik M., *Gromadzenie i przetwarzanie danych mających związek ze zwalczaniem przestępczości finansowej. Zasady dostępu, ograniczenia prawne*, WAT, Warszawa 2016.
- Kiedrowicz M., *Dostęp do publicznych zasobów danych – Big data czy Big brother*, w: *INTERNET. Publiczne bazy danych i Big data*, red. G. Szpor, s. 15–39, C.H. Beck, Warszawa 2014.
- Kiedrowicz M., *Location with the Use of the RFID and GPS Technologies – Opportunities and Threats*, GIS ODYSSEY 2016, Zagreb 2016.
- Kiedrowicz M., *Objects Identification in the Informations Models Used by Information Systems*, GIS ODYSSEY 2016, Zagreb 2016,
- Kiedrowicz M., *Organizacja i dostęp do heterogenicznych, publicznych zasobów danych*, w: *Projektowanie systemów informatycznych: modele i metody*, red. T. Nowicki, Z. Tarapata, WAT, Warszawa 2014.
- Kiedrowicz M., *Rejestry i zasoby informacyjne wykorzystywane przez organy odpowiedzialne za wykrywanie i przeciwdziałanie przestępczości*, w: *Jawność i jej ograniczenia: zadania i kompetencje*, red. G. Szpor, Wydawnictwo C.H. Beck, Warszawa 2015.
- Kiedrowicz M., *Rejestry publiczne wykorzystywane przez organy odpowiedzialne za wykrywanie i przeciwdziałanie przestępczości*, w: *Rejestry publiczne: jawność i interoperacyjność*, red. A. Gryszczyńska, C.H. Beck, Warszawa 2016.
- Kiedrowicz M., *Uogólniony model danych w rozproszonych rejestrach ewidencyjnych*, w: „Roczniki Kolegium Analiz Ekonomicznych” 2014, vol. 33.
- Kiedrowicz M., Koszela J., *Business Processes Modelling for the Processing of Classified Documents Using RFID Technology*, “Collegium of Economic Analysis Annals” 2016, vol. 42.

- Kiedrowicz M., Protasowicki T., Stanik J., *Wybrane aspekty standaryzacji w ochronie publicznych zasobów informacyjnych i świadczonych usług w kontekście społeczeństwa informacyjnego*, „Zeszyty Naukowe Uniwersytetu Szczecińskiego – Ekonomiczne Problemy Usług” 2014, z. 113.
- Kiedrowicz M., Stanik J., *Selected Aspects of Risk Management in Respect of Security of the Document Lifecycle Management System with Multiple Levels of Sensitivity*, w: *Information Management in Practice*, red. B.F. Kubiak, J. Maślankowski, University of Gdańsk, Gdańsk 2015.
- Koszela J., *Hurtownie procesów – systemy analitycznego przetwarzania procesów*, w: *Projektowanie systemów informatycznych: modele i metody*, red. T. Nowicki, Z. Tarapata, WAT, Warszawa 2014.
- Vukotic A., Watt N., Abedrabbo T., Fox D., Partne J., *Neo4J in action*, Manning Publications Co., Shelter Island USA 2015.
- Warchał Ł., *Using Neo4j graph database in social network analysis*, „Studia Informatica” 2012, vol. 33, no. 2A.
- Zarządzanie informacjami wrażliwymi. Wybrane aspekty organizacyjne, prawne i techniczne ochrony informacji niejawnych*, red. M. Kiedrowicz, WAT, Warszawa 2015.

## Źródła sieciowe

Neo4J, <https://neo4j.com/> (09.11.2016).

\* \* \*

## IAFEC – System Architecture

### Abstract

The paper has presented IAFEC system architecture (Information Analysis of Financial and Economic Crime), which enables the use of various data resources to combat economic crime (mainly in the area of finance). By creating the architecture of the system, the assumption was adopted that the required data (in particular, relating to the public registers and records) will be available without any formal and legal restrictions. The proposed architecture is presented from the point of view of the role of “an analyst” who is responsible for preparing data for analysis, to carry out this analysis and to present its results in the further proceedings. The size of the proposed use of resources and their variability makes the previously used tools and methods for data analysis insufficient.

**Keywords:** financial crime, money laundering, process modeling

