



BIG DATA JAKO ŹRÓDŁO INFORMACJI ROZSZERZAJĄCE FUNKCJONOWANIE SYSTEMÓW WSPOMAGANIA DECYZJI W ZARZĄDZANIU KRYZYSOWYM

Stanisław Drosio¹, Stanisław Stanek²

¹Uniwersytet Ekonomiczny w Katowicach
Wydział Informatyki i Komunikacji

²Wyższa Szkoła Oficerska Wojsk Lądowych im. gen. Tadeusza Kościuszki
Wydział Zarządzania

Streszczenie: W związku z bardzo dynamicznym rozwojem koncepcji i zastosowania Big Data w różnych obszarach aktywności człowieka coraz częściej mówi się o możliwości wykorzystania tego sposobu analizy danych w działaniach administracji publicznej. Z racji szerokiego spektrum analiz i przedstawienia koncepcji zastosowania hybrydowego podejścia we wspieraniu decyzji podczas realizacji procesów zarządzania kryzysowego autorzy prezentują wartość dodaną, jaką niesie wykorzystanie Big Data w analizach heterogenicznych źródeł danych, z jakich czerpie administracja kryzysowa. Artykuł został podzielony na trzy zasadnicze części. W pierwszej opisano genezę wykorzystania koncepcji Big Data w analizach decyzyjnych. W drugiej ukazano Big Data jako kolejne ogniwo ewolucji systemów informatycznych, szczególnie tzw. Business Intelligence, oraz pokazano, w jaki sposób koncepcja ta może wesprzeć Hybrydowy System Wspomagania Decyzji w Warunkach Kryzysu. W trzeciej części artykułu zaprezentowano kilka przykładów wykorzystania Big Data w szeroko rozumianym zarządzaniu bezpieczeństwem, a także przybliżono wszelkie konsekwencje, jakie niesie ze sobą wykorzystanie zbiorów Big Data w praktyce. W podsumowaniu ukazano dalsze plany rozwoju koncepcji oraz zgrupowano wnioski wynikające z przeprowadzonej analizy.

Słowa kluczowe: Big Data, Crisis Management, Decision Support System, Hybrydowy System Wspomagania Decyzji Zarządzania Kryzysowego

DOI: 10.17512/znpcz.2017.2.10

Wprowadzenie

Poczta elektroniczna, lokalizacja urządzeń przenośnych, szeroko rozumiane media społecznościowe, takie jak m.in.: Facebook, LinkedIn, Flickr, Twitter, coraz popularniejsze blogi oraz wszechobecna mobilność urządzeń i użytkowników powodują wykładniczy wzrost ilości danych, które można poddać analizie. Celem takich analiz może być koncentrowanie działań reklamowych, przydzielenie sprzedawców na potrzeby klientów, śledzenie ruchów w sieci czy przekazywanie do działu sprzedaży banku informacji na temat zmian środków na rachunku, a tym samym skierowanie sprzedawców z odpowiednimi produktami do klienta. Analizy te mogą dać również ogromną wartość dodaną dla organów administracji publicznej czuwających przez cały czas nad bezpieczeństwem obywateli, a pracę swoją opierających na słabych i silnych sygnałach pochodzących z różnych źródeł. Auto-

rzy niniejszego opracowania przyjęli sobie za cel przeprowadzenie analizy i scharakteryzowanie możliwości wykorzystania tych ogromnych zbiorów danych za pomocą bijącej od kilku lat rekordy popularności koncepcji opisywanej terminem Big Data. W artykule posłużono się szeregiem przykładów oraz sytuacji, w których Big Data stanowi wartość dodaną w walce o bezpieczeństwo i ciągłość działania gospodarki na danym terenie. Artykuł przedstawia również szerokie spektrum ryzyk, jakie niesienie ze sobą gromadzenie, przetwarzanie i udostępnianie danych, które wielokrotnie stanowią dane wrażliwe osób prywatnych lub przedsiębiorstw i mogą stać się bardzo ciekawym celem nie tylko dla oszustów czy pospolitych przestępców, ale również dla grup terrorystycznych, szczególnie w momencie, gdy wyniki analiz wskazują słabe punkty systemu bezpieczeństwa publicznego. Rozszerzając koncepcję Hybrydowego Systemu Wspomagania Decyzji (Stanek, Drosio 2012; Stanek, Namysło, Drosio 2013) o użycie Big Data, autorzy starali się przedstawić ogromny potencjał oraz nośność idei hybrydowości w przypadku tak turbulentnego i zmiennego środowiska decyzyjnego jak sytuacje kryzysowe czy brak ciągłości działania.

Big Data jako kolejny element na ścieżce rozwoju Business Intelligence

Podstawowym celem przyświecającym twórcom pierwszych systemów wspomagania decyzji, który pozostaje aktualny do dzisiaj, jest dostarczyć właściwą informację do właściwej osoby w odpowiednim formacie, miejscu i czasie (Stanek, Drosio 2012). Zwykle definicje powyższych środowisk informatycznych są bardzo mocno uogólnione. W niniejszej analizie została użyta definicja opracowana przez Kleina i Methliego: system wspomagania decyzji to system informatyczny, który dostarcza informacje w danej dziedzinie przy wykorzystaniu analitycznych modeli decyzyjnych z dostępem do baz danych w celu wspomagania decydentów w skutecznym działaniu kompleksowym i źle strukturalizowanym środowisku (Drosio 2010).

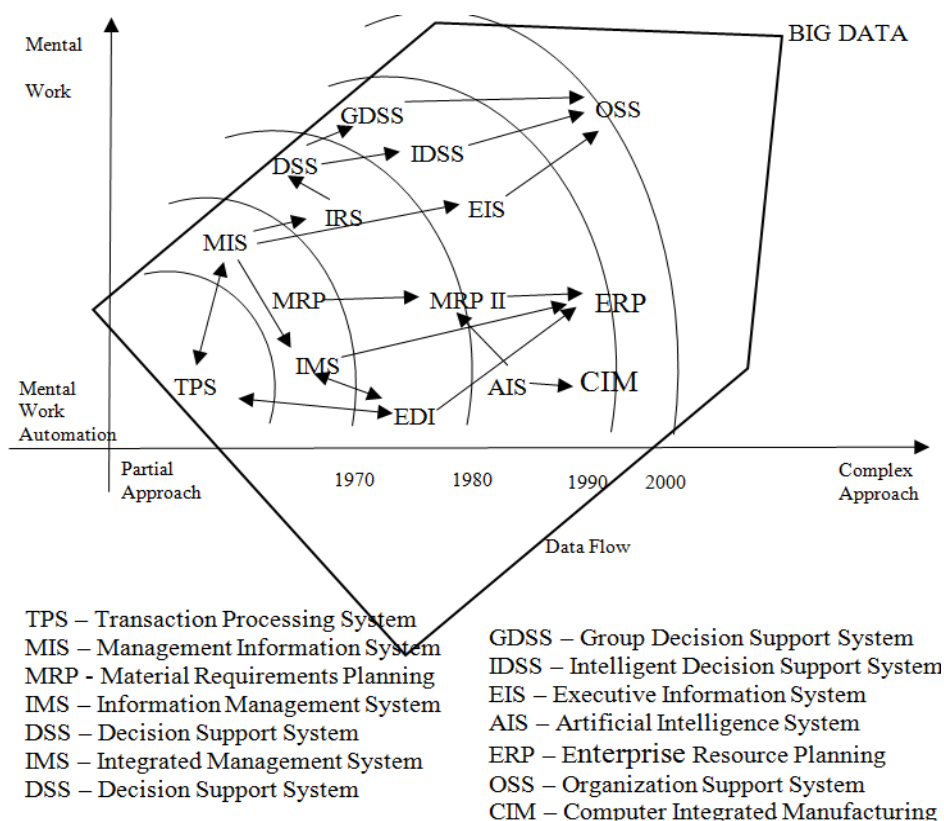
Ewolucja narzędzi IT w stronę tego typu systemów było skutkiem coraz większych potrzeb raportowych, jakie użytkownicy stawiali przed zintegrowanymi systemami zarządzania. Systemy informatyczne bazujące na transakcyjnej bazie danych nie były w stanie sprostać wszystkim tym potrzebom.

Data Mining a Big Data

Wskutek zmian opisanych powyżej narodziła się idea niezależnych systemów umożliwiających przekazywanie użytkownikowi w sprawny sposób informacji mogących wspierać proces decyzyjny, które jednak ze względu na ilość przetwarzanych danych nie zostały oparte na transakcyjnej bazie danych, ale na nowym typie zasobu, nazywanego hurtownią danych (Stanek 2007).

Jednakże zarówno hurtownie danych, jak również bazy transakcyjne niosły ze sobą potrzebę prowadzenia analiz w sposób sekwencyjny, często uniemożliwiający ze względu na charakter narzędzi ETL (*Extract, Transform and Load*) zasilających hurtownię danych prowadzenie analiz ad hoc, co w znacznym stopniu obniżało efektywność narzędzi. To właśnie na etapie eksploatacji powyżej opisanych

narzędzi oraz hurtowni danych pojawiła się subdyscyplina nauk informatyki i statystyki nazywana dzisiaj Data Mining. Odkrywanie wiedzy, jak często określany jest proces Data Mining, polega na wyszukiwaniu korelacji danych pod kątem ich powiązań oraz zależności, a narzędzia do tego wykorzystywane w głównej mierze bazują na paradygmacie hurtowni danych. Struktura ta przechowuje dane historyczne oraz obrazuje pewien dynamizm zmian za pomocą odkładania kolejnych „warstw danych”. Nie umożliwia prowadzenia analizy on-line na zmieniających się w sposób ciągły danych wejściowych, ponieważ zależna jest od częstotliwości i jakości zasileń (Wieczorkowski 2014).



Rysunek 1. Ewolucja systemów informatycznych wspomagających zarządzanie

Źródło: (Stanek 2007)

Aby bardziej obrazowo przedstawić umiejscowienie koncepcji Big Data wśród innych koncepcji gromadzenia i analiz oraz wizualizacji danych w systemach informatycznych, na *Rysunku 1* przedstawiono uproszczoną historię ewolucji systemów IT w zarządzaniu. Proces przedstawiony na powyższym rysunku jest ciągły, jednakże na osi czasu co jakiś czas pojawiają się nowe koncepcje zastosowania wsparcia informatyki w działaniu podmiotów gospodarczych. Próba umiejscowienia Big Data jako jednego z wymienionych koncepcji jest praktycznie

niemożliwa, dlatego na powyższym rysunku ukazany został za pomocą strzałki kierunek symbolizujący przepływ danych do Big Data, które obecnie stanowią niejako „pokrywkę” wszelkich źródeł danych, a przedstawione na rysunku systemy, są jedynie ich podzbiorem. Schematyczny zbiór źródeł danych, z których czerpie koncepcja Big Data, został przedstawiony na *Rysunku 1*. Maksymalnie uogólniając analizę zbioru informacji znajdujących się na wspomnianej wizualizacji, możemy określić Big Data jako dalszy etap ewolucji rozwoju przetwarzania analitycznego zaliczanego w literaturze do grona rozwiązań Business Intelligence (Wieczorkowski 2014) wspierających proces przekształcania danych w informacje użyteczne dla decydentów.

Przejście od tradycyjnego przetwarzania analitycznego w kierunku analityki Big Data to rozległa transformacja, obejmująca między innymi:

- Transformację od tradycyjnej hurtowni danych w kierunku logicznej hurtowni danych. Serwer wirtualizacji (np. Denodo Platform) ukrywa rzeczywiste umiejscowienie danych (np. serwer bazy SQL, system NoSQL, plik Hadoop, arkusz Excela). Zanika potrzeba ich przenoszenia (ETL), utrzymywania rozbudowanej fizycznej infrastruktury, dane są dostępne na bieżąco, zmniejsza się zakres prac, czas budowy i implementacji oraz ryzyko związane z wdrażaniem hurtowni danych.
- Transformację metod przetwarzania analitycznego w kierunku metod ukierunkowanych na przetwarzanie równoległe, takich jak algorytm MapReduce, Hadoop Distributed File System. Preferowane są rozwiązania dobrze pracujące w gridach, chmurze obliczeniowej.
- Transformację tradycyjnych narzędzi przetwarzania analitycznego w kierunku rozwiązań wspomagających Big Data. Dla przykładu SQL Server 2016 integruje się z Hadoop (np. poprzez zapytania T-SQL wykonywane z wykorzystaniem PolyBase), dostarcza wbudowane wspomaganie dla formatu JSON, implementuje język R.
- Transformację tradycyjnych narzędzi wizualizacji dostarczających nowych rozwiązań dla obrazowania dynamicznie zmieniających się danych. Przykładem może być wykorzystanie ArcGIS w zastosowaniu do aktualnych stosowanych przez przestępców taktyk prania brudnych pieniędzy. Jedną z taktyk jest rozbijanie dużych kwot na małe przelewy. Wykorzystanie funkcji *join features* (cech łączonych) umożliwiło wyodrębnienie miliardów transakcji, których przetwarzanie w klastrze 140 maszyn zajęło w prezentacji firmy ESRI 48 minut, dając w wyniku obrazowanie dynamiki strumieni dopływających do tego samego odbiorcy (narzędzie GeoAnalytics).
- Transformację zadań tradycyjnego analityka danych w kierunku szerokiego spojrzenia z perspektywy pracownika wiedzy. Pracownik wiedzy powinien mieć możliwość zbierania wymagań i danych, ich analizy poprzez wizualizację i modelowanie. Potrzebne są umiejętności w zakresie analiz statystycznych, BI, Big Data, BPM oraz ACM (Van der Alst, Berens 2001; Osuszek, Stanek, Twardowski 2016).

Big Data jako rozszerzenie Hybrydowego Systemu Wspomagania Decyzji

W nawiązaniu do przedstawionej kilka lat temu koncepcji Hybrydowego Systemu Wspomagania Decyzji w Warunkach Zagrożenia (Stanek, Drosio 2012) można nakreślić ciekawy ze względów badawczych oraz posiadający ogromny potencjał predykcyjno-analityczny obszar wykorzystania koncepcji wielkich zbiorów danych. Dlaczego w przypadku zarządzania kryzysowego oraz utrzymania ciągłości działań jest to wyłącznie narzędzie wspomagające, przedstawiają autorzy w dalszej części niniejszej analizy.

W tym miejscu należy zastanowić się, w jaki sposób możemy wykorzystać potencjał tego narzędzia w kontekście wyżej wymienionego obszaru wspomagania decyzji zarządczych. W Tabeli 1 przedstawiona została charakterystyka poszczególnych modułów wchodzących w skład koncepcji autorów, wraz z potencjalnie wykorzystywaną technologią wspierającą dany obszar działania.

Tabela 1. Krótka charakterystyka poszczególnych elementów koncepcji Platformy „Czterech Światel”

Platforma (moduł)	Obszar odpowiedzialności – wspomaganie	Proponowane do wykorzystania technologie
Niebieskie Światło	Planowanie, w tym utrzymanie i rozwój planów kryzysowych, odbudowa, zapobieganie i sprawozdawczość, informowanie/rozwój/szkolenie społeczności lokalnej oraz personelu, w tym wsparcie psychologiczne, kształtowanie wizerunku, w tym relacje z mediami, pomoc poszkodowanym.	Spersonalizowany portal informacji i wiedzy, e-video learning, virtual reality, symulacje, e-learning, cloud sourcing, WEB 2.0, 3.0, portale społecznościowe, gry i zabawy, badania społeczne.
Zielone Światło	Monitoring / zintegrowane wczesne ostrzeżenie na podstawie lokalnego, krajowego, międzynarodowego systemu czujników (np. Pogodynka), łączenie (<i>fusing</i>) informacji z systemów wczesnego ostrzegania. Zapobieganie poprzez przekazywanie informacji do pozostałych platform.	Hurtownia danych, system ekspertowy, szyna danych, ontologia dziedzinowa, system reguł rozmytych do wykrywania zagrożeń, Hadoop oraz inne technologie Big Data, analityki.
Żółte Światło	Weryfikacja, szczegółowa ocena stanu zagrożenia. Wybór oraz dostosowanie planu reagowania. Ewentualne powołanie oraz obsługa sztabu kryzysowego, zespołów roboczych, szacowanie, aplikowanie oraz obsługa/rozdział niezbędnych zewnętrznych zasobów. Koordynacja i nadzór dalszych działań kryzysowych wszystkich służb, zespołów oraz wolontariuszy.	Wideokonferencje, serwery integracyjne, modele optymalizacji wielokryterialnej, zaawansowane narzędzia wyszukiwania informacji, ACM, kokpit menedżerski, systemy wieloagentowe, GIS, EIS.

Czerwone Światło	Platforma wspierająca bezpośrednio działania służb ratunkowych oraz umożliwiająca raportowanie niezbędne do realizacji fazy odbudowy i odtworzenia infrastruktury oraz zabezpieczenia uszkodzonych. Wsparcie bezpośrednio oparte na systemie mapowym oraz technologiach umożliwiających bieżące współdziałanie służb oraz nadzór nad terenami zagrożonymi kryzysem, segregacja uszkodzonych, przemieszczanie ludzi i dobytku, logistyka i zaopatrzenie, zapewnienie porządku i bezpieczeństwa, w szczególności zabezpieczenie terenu, niedopuszczenie do paniki, kradzieży, minimalizacja strat), raportowanie i obrazowanie, uporządkowana wielowariantowość, elastyczność i niezawodność działań, solidna sygnalizacja i sprawozdawczość.	GIS, systemy dronów przekazujących obraz on-line, terminale GPS, system wideokonferencyjny, reporting services, systemy rezerwowe i krytyczne systemy autonomiczne.
------------------	---	---

Źródło: (Stanek, Drosio 2012)

Głównymi zadaniami stawianymi przed tą koncepcją są (Stanek, Namysło, Drosio 2013): (1) integrowanie heterogenicznych danych za pomocą odpowiednio przygotowanej ontologii dziedzinowej, (2) przetwarzanie danych w celu jej udostępnienia jako pomoc zarówno dla użytkowników front – end (mieszkańców obszaru administracji wykorzystującego narzędzie), jak i back – end (organów administracji i przedstawicieli biznesu włączonych w system antykryzysowy), (3) prowadzenie bieżącej analizy w oparciu o innowacyjne narzędzia i metodyki (logika rozmyta, agenci oprogramowania, modelowanie semantyczne), (4) wspieranie decydentów odpowiednimi danymi dostarczonymi w odpowiedniej formie i czasie, w celu ułatwienia celowych i efektywnych decyzji na podstawie dostępnych danych.

Analizując powyższe, można założyć, że Big Data wpisuje się w koncepcję Platformy „Zielonego Światła”, gdzie podstawowym założeniem systemu jest wspieranie poprzez uzyskiwanie możliwe szerokiego zastosowania różnorodnych danych maksymalnie dużych zasobów danych, które następnie trzeba poddać odpowiedniej analizie w celu wnioskowania mogącego w praktyce przyczynić się do podniesienia poziomu bezpieczeństwa na danym obszarze odpowiedzialności.

Kryzysowy proces decyzyjny w kontekście Big Data

Na podstawie polskiego prawodawstwa proces zarządzania kryzysowego określa się jako działalność administracji publicznej będąca elementem kierowania bezpieczeństwem narodowym, która polega na zapobieganiu sytuacjom kryzysowym, przygotowaniu do przejmowania nad nią kontroli w drodze zaplanowanych działań, reagowania w przypadku wystąpienia sytuacji kryzysowej oraz na odtworzeniu infrastruktury lub przywróceniu jej pierwotnego charakteru (Ustawa z dnia 27 kwietnia 2007 r., art. 2). Jest to definicja bardzo ogólna, mówiąca o za-

rządaniu kry-zysowym w momencie pojawienia się kryzysu zagrażającego mieszkańcom znajdującym się na terenie odpowiedzialności danej jednostki samorządu terytorialnego. W związku z tym na potrzeby tego artykułu wykorzystywane jest szersze spojrzenie, bazujące na określonej w innych aktach prawnych oraz regulacjach międzynarodowych bazie stanów funkcjonowania systemu zarządzania kryzysowego oraz utrzymania ciągłości działań, które oparte jest na modelu trójeta-powym stosowanym głównie w Europie i Stanach Zjednoczonych Ameryki Północnej (Stanek, Drosio 2012):

1. faza prewencji i ostrzegania,
2. faza kryzysu,
3. faza usuwania skutków kryzysu.

W kontekście tematyki niniejszego artykułu oraz założeń, jakie przyświecają orędownikom Big Data, przydatna może być wiedza czerpana z tych zbiorów danych właśnie do utrzymania odpowiedniej jakości i skuteczności pierwszej fazy procesu zarządzania kryzysowego, a więc prewencji i ostrzegania. Dzięki odpowiedniemu nasyceniu danymi pochodzącymi z różnych źródeł odpowiednio przygotowane narzędzia mogą wspomagać przynajmniej podstawowe analizy dające wstępne diagnozy i hipotezy. Wyniki te, poddawane następnie dyskusji i analizie w gronie osób odpowiedzialnych za bezpieczeństwo oraz ciągłość działań, mogą prowadzić, z odpowiednim wyprzedzeniem, do odnajdywania oraz neutralizowania syndromów i pierwszych objawów kryzysu lub nieciągłości.

Podstawowe charakterystyki Big Data

Wyróżniamy dwie podstawowe grupy Big Data (Tamhane, Sayyad 2014):

- Dane ustrukturalizowane – liczby i słowa, które mogą być w łatwy sposób kategoryzowane i analizowane. Te dane zazwyczaj pochodzą z sensorów urządzeń elektronicznych, takich jak smartfony, GPS czy z innych urządzeń mobilnych. W zbiór danych ustrukturalizowanych włączamy również takie elementy jak dane sprzedażowe, bilanse kont oraz parametry transakcji.

Ta grupa danych jest powszechnie wykorzystywana zarówno w analizach Business Intelligence, jak również w zarządzaniu kryzysowym, gdzie tego typu dane pochodzą z różnorodnych transakcyjnych systemów operacyjnych służb, administracji publicznej oraz raportów pochodzących z realizacji procedur i planów zarządzania kryzysowego. W przedstawianej przez autorów koncepcji Platformy „Czterech Świąteł” dane te stanowią fundament do budowy hurtowni danych, która stanowi podstawę dla modułu „Zielonego Światła” opisanego w *Tabeli 1*. Ten obszar danych stanowi obecnie podstawę do konstrukcji i eksploatacji systemów klasy Business Intelligence w obszarze zarządzania kryzysowego. Nowość w predykcji zdarzeń kryzysowych oraz nieciągłości działań stanowią analizy drugiej grupy danych określanych jako:

- Dane nieustrukturalizowane – zawierają bardziej złożone i kompleksowe dane niż dane ustrukturalizowane, takie jak opinie użytkowników na temat towarów i usług wprowadzone w sieci Internet, zdjęcia i inne multimedia oraz zawartość

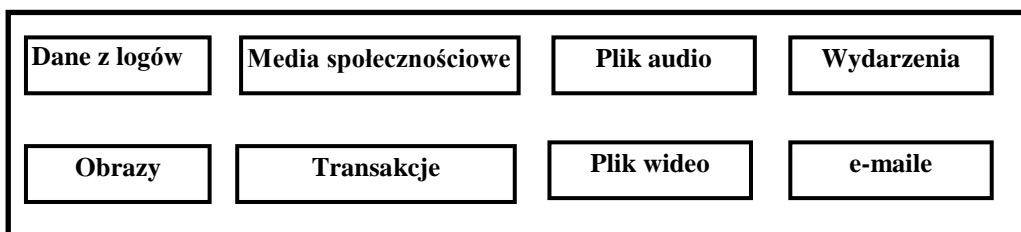
stron mediów społecznościowych. Te dane nie dają się łatwo kategoryzować i przeszukiwać.

Nieustrukturalizowane dane obszaru Big Data to po prostu rzeczy, które ludzie mówią – twierdzi Tony Jewitt Vice Prezes w Plano Consulting z Texasu (Thakur, Mann 2014). Powyższa grupa danych przede wszystkim charakteryzuje się bardzo małą formalizacją języka opisującego poszczególne elementy zbioru danych. Powoduje to duże problemy w maszynowym zadawaniu pytań, na które ma odpowiedzieć system analizujący bazę danych, w której te dane się znajdują. To właśnie ta grupa stanowi wyzwanie dla koncepcji wsparcia procesu decyzyjnego zarządzania kryzysowego opisanego powyżej, a zarazem jest niesamowitym rezerwuarem danych, na podstawie których można prowadzić predykcję pojawienia się zagrożeń lub kryzysów. W dalszej części niniejszego artykułu autorzy przytoczą kilka przykładów trafnych modeli predykcyjnych bazujących na nieustrukturalizowanych danych Big Data.

Aby bardziej obrazowo zademonstrować, z czym mamy do czynienia, warto tutaj przytoczyć kilka statystyk przedstawionych przez Usama Fayyada w czasie KDD BigMine 12 Workshop (Tamhane, Sayyad 2014). Obrazują one dobitnie, z jak ogromnym zbiorem danych mamy do czynienia:

- Każdego dnia Google przetwarza 1 miliard zapytań.
- Każdego dnia Twitter ma ponad 250 milionów wpisów.
- Każdego dnia Facebook ma ponad 800 milionów wpisów.
- Każdego dnia YouTube ma 250 miliardów wyświetleń.
- Dane produkowane przez ludzkość dzisiaj są estymowane w zettabajtach i każdego roku ich ilość przyrasta o 40%.
- Ogromne źródło danych to urządzenia mobilne oraz wielkie korporacje ICT (Google, Apple, Yahoo itp.).

Jak widać na powyższym zestawieniu, w przypadku Big Data trudno jest określić zamknięty zbiór danych, szczególnie w odniesieniu do drugiej grupy danych opisanej powyżej. W celu ukazania ogromnego ich zasobu na *Rysunku 2* przedstawione zostały wybrane źródła, z których mogą pochodzić dane określane jako Big Data. W celu ukazania ogromnego ich zasobu na *Rysunku 2* przedstawione zostały wybrane źródła, z których mogą pochodzić dane określane jako Big Data.



Rysunek 2. Przykładowe źródła danych Big Data

Źródło: (Wieczorkowski 2014)

Przykłady wykorzystania Big Data we wsparciu procesów zarządzania kryzysowego

W celu przeniesienia powyższych charakterystyk do obszaru zarządzania kryzysowego należy przytoczyć kilka podstawowych przykładów wykorzystanych obecnie przez służby bezpieczeństwa w celu skuteczniejszego zapewnienia bezpieczeństwa czy predykcji zdarzeń kryzysowych:

1. Przykłady danych ustrukturalizowanych: ilość opadów atmosferycznych, poziom wód w rzekach, natężenie ruchu samochodowego, ilość zdarzeń miejscowych w danej okolicy (np.: pożary, wypadki samochodowe), siła i kierunek wiatru, dane z osobistych odbiorników GPS osób poszkodowanych, natężenie ruchu z inteligentnych systemów zarządzania miastem, ilość zużytych sił i środków do likwidacji danego typu zdarzenia miejscowego (np.: ilość zużytego środka gaśniczego do ugaszenia pożaru 2-kondygnacyjnego domu).
2. Przykłady danych nieustrukturalizowanych: zdjęcia zamieszczone na profilach osobistych mediów społecznościowych rejonu zainteresowania, komentarze przed, w trakcie i po zdarzeniu kryzysowym na portalach społecznościowych w miejscu zdarzenia, nagrania z monitoringu miejskiego, komentarze pojawiające się w mediach społecznościowych w miejscu potencjalnego kryzysu (np.: informacje o braku prądu, braku dostępu do Internetu, komentarze o dziwnym zachowaniu ludzi lub zaistnieniu zjawisk dotychczas nie pojawiających się w tej okolicy).

Zdjęcia, wideo i media społecznościowe w zarządzaniu kryzysowym na przykładzie danych nieustrukturalizowanych

Jak zostało to już opisane powyżej, drugą grupą danych kwalifikowanych do zakresu Big Data są elementy danych nieustrukturalizowanych. Obecnie największe ich źródło stanowią wszelkiego rodzaju dane pochodzące z obszaru mediów społecznościowych, a więc zdjęcia, filmy, komentarze. Oczywiście wymienione tutaj źródła generują cały szereg danych ustrukturalizowanych, które mogą zostać wykorzystane z powodzeniem w sposób opisany powyżej.

W niniejszym przykładzie autorzy przedstawiają potencjalne i zaistniałe sytuacje umożliwiające wykorzystanie nieustrukturalizowanych Big Data w czasie działań powiązanych z ciągłością zarządzania i reakcjami kryzysowymi. Przytoczony przykład opiera się na grupie największej ilości nieustrukturalizowanych danych, jakie są powszechnie dostępne i często chętnie publikowane przez autorów lub osoby na nich się znajdujące, mowa oczywiście o zdjęciach. Korzystając z różnych metod nieparametrycznych, badamy „ten świat” za pomocą dużego zbioru obrazów zebranych z Internetu (Torralba, Fergus, Freeman 2008). Dodatkowo obecnie trend w rozwoju aparatów oraz innych rejestratorów obrazu zmierza w kierunku zapisu, nawet bez zgody i wiedzy twórcy zdjęcia, śladu GPS umożliwiającego zarejestrowanie szczegółowych danych dotyczących miejsca oraz czasu rejestracji zdjęcia (Datcu, Rothkrantz 2007).

Bazując na powyższych elementach, można przygotować następujący problem do rozwiązania na podstawie danych nieustrukturalizowanych w zbiorze Big Data:

1. Problem: w trakcie realizacji procedur kryzysowych zaszła konieczność zlokalizowania człowieka bądź grupy ludzi.

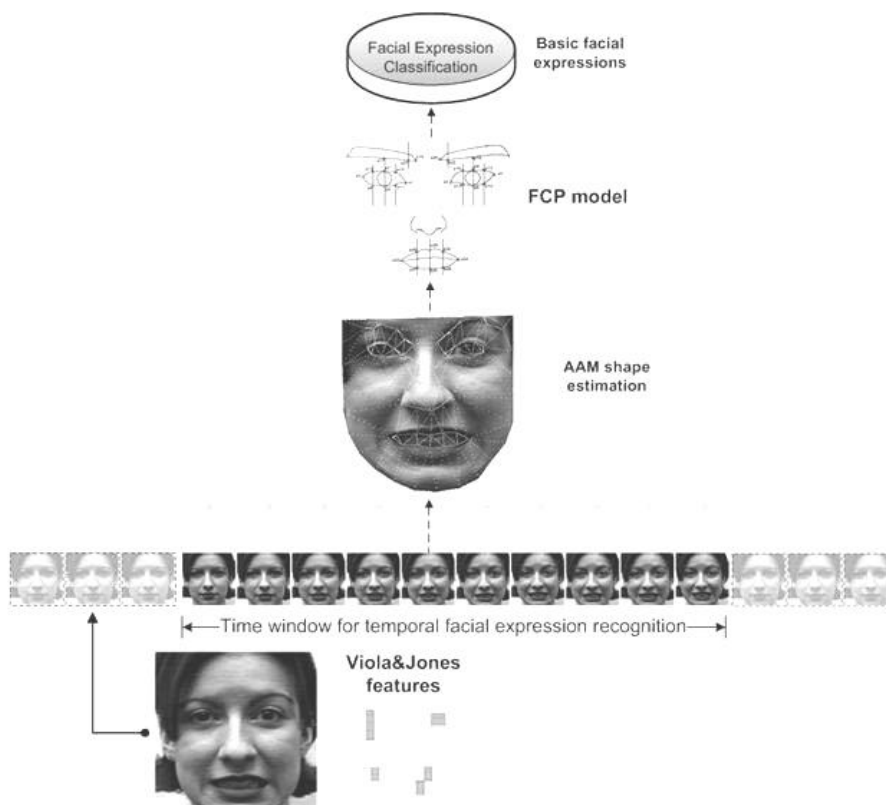
2. Lokalizacja problemu:

- Na podstawie danych pochodzących z monitoringu miejskiego jesteśmy w stanie, używając odpowiedniego oprogramowania, przeprowadzić analizę wykonanych na podstawie filmów wideo zdjęć twarzy przechodniów.
- W celu podniesienia efektywności pracy drugim źródłem danych zostały obrane media społecznościowe, a konkretnie zdjęcia umieszczane przez użytkowników portalu społecznościowego w wybranym obszarze poszukiwań.

Zbiór powyższych nieustrukturalizowanych danych został poddany obróbce narzędziem Viola & Jones Face Detector. Powyższe zastosowanie narzędzi do obróbki zdjęć jest pierwszym etapem w tego typu działaniach. Trwają już z sukcesem prowadzone eksperymenty umożliwiające nie tylko wykrywanie danej opisanej pewnym zakresem danych twarzy, ale również analizom poddawane są emocje ludzi zgromadzonych na określonej fotografii (*Rysunek 3*).

3. Efekt zastosowania analiz Big Data:

- Ułatwienie określania poprawności wyznaczenia obszaru poszukiwań zaginionych, określanie miejsca lub miejsc, w jakich mogła znajdować się poszukiwana osoba lub osoby.



Rysunek 3. Proces rozpoznawania twarzy w nawiązaniu do zdjęcia w systemie Viola & Jones

Źródło: (Datcu, Rothkrantz 2007)

Na podstawie powyższego przykładu jednoznacznie możemy wskazać, iż w związku z powszechnością oraz ilością dostępnych narzędzi do analiz obecnie najbardziej rozpowszechnionym typem danych wykorzystywanych w procesach zarówno ratownictwa, jak i szeroko rozumianego zarządzania kryzysowego są dane ustrukturalizowane wchodzące w skład Big Data. W związku z coraz większą dostępnością narzędzi oraz rozwojem technologii sprawne i szybkie analizy danych, jak przedstawione powyżej, staną się coraz bardziej dostępne w codziennej pracy struktur zarządzania kryzysowego.

Analiza SWOT wykorzystania Big Data we wspieraniu procesów zarządzania kryzysowego

Należy jednakże pamiętać, iż operowanie na stale przyrastających zbiorach danych, które w wielu przypadkach posiadają tzw. Data Reservoirs (Raviendran 2014), niesie ze sobą cały szereg zagrożeń i możliwych do materializacji ryzyk. W celu przybliżenia tej materii autorzy pokusili się o przygotowanie za pomocą jednej z najbardziej rozpowszechnionych w pracach sektora ICT metody – analizy SWOT. W Tabeli 2 zaprezentowano zestawienie najważniejszych wewnętrznych i zewnętrznych czynników przemawiających za, jak również i przeciwko powszechnemu wykorzystaniu Big Data jako narzędzia wspierającego podejmowanie decyzji w procesach zarządzania kryzysowego.

Z analizy poniższego zestawienia koncepcji Big Data wyłania się bardzo niejednoznaczny obraz. Szczególnie dotyczy to ich wykorzystania przez instytucje rządowe, które z jednej strony stoją na straży bezpieczeństwa, a z drugiej coraz częściej podejrzewane są o inwigilację obywateli. Niestety ten drugi obraz potwierdzają takie przykłady jak ujawnione przez Edwarda Snowdena, byłego współpracownika National Security Agency. To z rewelacji przekazanych przez tego człowieka dowiedzieliśmy się o powszechnym wykorzystaniu przez służby takich narzędzi, jak: XKeyScore, Tempora, Bullrun czy EdgeHill, które pozwalają na śledzenie na bieżąco aktywności „celów” działania służb. Takie działania są możliwe dzięki szerokiej analizie elementów określanych jako zbiory Big Data: e-maili, rozmów prowadzonych za pomocą komunikatorów, aktywności w serwisach społecznościowych i wyszukiwarkach internetowych. Informacje Snowdena ujawniły również nieskuteczność szyfrowania komunikacji elektronicznej w stosunku do działań służb (Żmudziński 2014).

Tabela 2. Analiza SWOT wykorzystania Big Data we wsparciu procesów zarządzania kryzysowego

	Pozytywne	Negatywne
Wewnętrzne	<p>Mocne strony:</p> <ul style="list-style-type: none"> - Potencjał umożliwiający zwiększenie zaangażowania w prewencję, a nie wyłącznie odpowiedź na pojawienie się kryzysu. - Praktycznie nieskończony zasób danych i informacji. - Możliwość wykorzystania takich elementów, jak: <i>crowdsourcing</i>, czy <i>customer to administration</i>, celem podniesienia jakości działań. - Realizowanie podstawowego założenia, jakim jest maksymalizacja liczby źródeł informacji umożliwiających przewidywanie i przeciwdziałanie kryzysom. 	<p>Słabe strony:</p> <ul style="list-style-type: none"> - Ciągły przyrost ilości danych powoduje ogromny popyt na sprzęt umożliwiający katalogowanie i analizy coraz większych zbiorów. - W przypadku Unstructured Data pojawia się problem z przygotowaniem odpowiednich założeń do ich analizy i przetwarzania, co powoduje potrzebę rozpoczęcia prac od odpowiednich projektów R&D. - Prowadzenie analiz tego typu zbiorów danych narażone jest na szereg problemów opisanych w HACE theorem, co może powodować błędne wnioski i co za tym idzie –niemożność wykorzystania efektów pracy w praktyce.
Zewnętrzne	<p>Szanse:</p> <ul style="list-style-type: none"> - Duży potencjał rozwoju oraz możliwości wyszukiwania i uruchamiania coraz to nowych koncepcji analiz. - Stosunkowo nowy trend, co oznacza brak skostniałych i zastanych narzędzi i sposobów pracy w tym obszarze. - Komplementarność z hurtowniami danych oraz Data Mining powodują ewolucyjność, a nie rewolucyjność procesu wprowadzania. - Brak jasno określonych składowych Big Data daje praktycznie nieograniczone możliwości rozwoju w kierunku włączania do zbioru kolejnych rodzajów danych. 	<p>Zagrożenia:</p> <ul style="list-style-type: none"> - Ze względu na wiele źródeł danych Big Data niezgromadzone w centralnej bazie na potrzeby zarządzania kryzysowego są bezużyteczne w przypadku nieciągłości takich elementów infrastruktury, jak sieci komórkowe oraz Internet. - Wykorzystanie rozległych (rosnących w sposób ciągły) i rozproszonych baz danych może uniemożliwić wykorzystanie elementów i procedur Data Mining, chociażby ze względu na koszty przenoszenia i magazynowania danych. - Gromadzenie i przetwarzanie danych wrażliwych stanowiących dużą część Big Data powoduje wiele wątpliwości natury podstawowej i wpływa na negatywne postrzeganie takich zbiorów danych jako źródła inwigilacji społeczeństwa. - W nawiązaniu do powyższego wiele krajów stara się regulować prawnie dostęp i przechowywanie takich zbiorów, co może prowadzić do przeregulowania i bezużyteczności takich szczątkowych informacji.

Źródło: (Drosio, Stanek 2016)

Podsumowanie

Statystyki pokazują, iż spośród 40 zettabajtów danych, jakie mają być zgromadzone przez ludzkość do 2020 roku, jak to określa IDC, analizom podlega zaledwie 0,5% (Żmudziński 2014). Takie statystyki wydają się nastrajać optymistycznie i pokazują, iż zainteresowanie wszelkimi danymi pochodzącymi od poszczególnych obywateli nie jest aż tak duże. Co za tym idzie – nieuprawniona wydaje się być na dzień dzisiejszy obawa przed powszechną inwigilacją i tworzeniem na podstawie Big Data swoistego Big Brothera. Natomiast pokazuje to, jak ogromny potencjał daje wykorzystanie tych danych nie tylko do celów komercyjnych, ale również do ratowania zdrowia i życia ludzkiego w ramach działań struktur zarządzania kryzysowego. W związku z tym zmiana czasów i dostępnej technologii powinna nieść ze sobą jeszcze jedną zmianę, która zdaje się być najtrudniejsza do realizacji. Decydenci i struktury kryzysowe muszą zmienić sposób podejścia do swojej misji.

Konkludując, należy spojrzeć na zarządzanie kryzysowe nie jak na zbiór procedur i technik, jakie należy zastosować w przypadku pojawienia się kryzysu, ale całą siłę i możliwości działania należy skierować na przewidywanie i przeciwdziałanie zagrożeniom, których materializacja powoduje kryzys (Bertrand, Chris 2002). W nawiązaniu do analizy przeprowadzonej powyżej wykorzystanie Big Data wydaje się być jednym z elementów umożliwiających zmianę podejścia, dzięki ogromnemu potencjałowi oraz ciągle nie do końca określonym możliwościom, jakie idą za zastosowaniem zaawansowanych analiz danych.

Literatura

1. Bertrand R., Chris L. (2002), *A New Approach to Crisis Management*, „Journal of Contingencies and Crisis Management”, Vol. 10, Issue 4, s. 181-191.
2. Datcu D., Rothkrantz L.J.M. (2007), *The Use of Active Appearance Model for Facial Expression Recognition in Crisis Environments*, [w:] Van de Walle B., Burghardt P., Nieuwenhuis C. (eds.), *Proceedings ISCRAM 2007*, Conference Paper, Netherlands.
3. Drosio S. (2010), *Computer Support in Decision Making of Live-Saving System*, Promotor, Warszawa.
4. Drosio S., Stanek S. (2016), *The Big Data Concept As a Contributor of Added Value to Crisis Decision Support Systems*, [w:] Sammon D. i in. (eds.), *Proceedings of the 2016 Open Conference of the IFIP WG 8.3, “Big Data, Better Decisions, Brighter Future”*, University College Cork, Cork, s. 228-239.
5. Osuszek L., Stanek S., Twardowski Z. (2016), *Leverage Big Data Analytics for Dynamic Informed Decisions with Advanced Case Management*, „Journal of Decision Systems”, Vol. 25, No. S1, s. 436-449.
6. Raviendran M. (2014), *Big Data Reservoirs: Getting from Big Data to Valuable Data*, Oracle, https://blogs.oracle.com/dataintegration/entry/big_data_reservoirs_getting_from (dostęp: 10.11.2015).
7. Stanek S., Drosio S. (2012), *A Hybrid Decision Support System for Disaster/Crisis Management*, [w:] Respicio A., Burstein F. (eds.), *Fusing Decision Support into the Fabric of the Context*, IOS Press, Amsterdam, s. 279-290.
8. Stanek S. (2007), *Metodologia budowy komputerowych systemów wspomagania organizacji*, Akademia Ekonomiczna im. Karola Adamieckiego w Katowicach, Katowice.

9. Stanek S., Namysło J., Drosio S. (2013), *Developing the Functionality of a Mobile Decision Support System*, „Journal of Decision Systems: Special Issue on Mobile Decision Support Systems: Addressing Challenges of Real-Time Decision-Making”, Vol. 22, No. 1, s. 53-68.
10. Tamhane D., Sayyad S. (2014), *Big Data Analysis Using HACE Theorem*, „International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)”, Vol. 4, Issue 1, s. 18-23.
11. Torralba A., Fergus R., Freeman W.T. (2008), *80 Million Tiny Images: A Large Data Set for Nonparametric Object and Scene Recognition in Pattern Analysis and Machine Intelligence*, „IEEE Transactions on Pattern Analysis and Machine Intelligence”, Vol. 30, No. 11, s. 1958-1970.
12. Ustawa z dnia 27 kwietnia 2007 r. o zarządzaniu kryzysowym (tekst jedn.: Dz.U. 2017 poz. 209).
13. Van der Alst V.M.P., Berners P.J.S. (2001), *Beyond Workflow Management: Product – Driven Case Handling*, [w:] Ellis S., Rodden T., Zigurs I. (eds.), *International ACM SIGGROUP Conference on Supporting Group Work (GROUP 2001)*, ACM Press, New York, s. 42-51.
14. Wieczorkowski J. (2014), *The Use of Big Data Concept in Public Administration*, [w:] *Collegium of Economic Analysis Annals*, Örebro University School of Business, Örebro, s. 567-579.
15. Żmudziński J. (2014), *Big Data – szanse i zagrożenia dla bezpieczeństwa*, Biuletyn PTI, <http://www.biuletyn.pti.org.pl/biuletyn-pti/bezpieczenstwo-informacji/317-bigdata-szanse-i-zagrozenia-dla-bezpieczenstwa> (dostęp: 10.11.2015).

BIG DATA AS A SOURCE OF EXTERNAL INFORMATION FOR CRISIS DECISION SUPPORT SYSTEM

Abstract: The paper seeks to emphasize the added value that stems from the use of Big Data in handling heterogeneous data sources accessed by crisis management structures. It is broken down into three main sections. The first one outlines the origins of the Big Data concept as an approach to data analysis and its position in the history of Decision Support Systems. The second part of the paper highlights the ways in which the concept could support the authors' own creation known as Hybrid Decision Support System for Crisis/Disaster Management. The third part of the paper uses an example to showcase applications of Big Data in security management at large and to illustrate the implications of using Big Data in crisis management practice. The conclusion brings together critical observations and judgments voiced in the paper supported by SWOT analysis of Big Data usage in Crisis Management support.

Keywords: Decision Support System, Big Data, Crisis Management, Hybrid Decision Support System for Crisis/Disaster Management