

Rozdział 1

Wprowadzenie do systemu DISESOR

Wojciech Moczulski, Piotr Przystałka, Marek Sikora, Dominik Ślęzak

1.1 Wstęp

W nowoczesnym górnictwie podziemnym powszechnie stosowane są systemy monitorowania maszyn oraz oceny zagrożeń. Współpracują one z układami czujników rozmieszczonych w podziemiach kopalń, a także umieszczonych na poszczególnych maszynach i urządzeniach. Systemy monitorowania i systemy dyspozytorskie należą zazwyczaj do klasy SCADA (Supervisory Control And Data Acquisition), MES (Manufacturing Execution Systems) lub AM/FMGIS (Automated Mapping/Facilities Management). Umożliwiają one gromadzenie danych, którymi są parametry procesowe, sygnały sterowania, sygnały wyjściowe itd., a także komunikatów generowanych przez systemy sterowania i zabezpieczeń, oraz przez obsługę. Dane te mają zazwyczaj charakter strumienia danych, gdyż napływają w sposób ciągły z wielu źródeł i z różną częstotliwością.

W obszarze górnictwa węgla kamiennego (dotyczy to również innych surowców) mamy do czynienia z bardzo dużym zróżnicowaniem systemów. Dostawcami najpopularniejszych systemów związanych z monitorowaniem zagrożeń naturalnych w górnictwie węgla kamiennego (np. metan, sejsmologia, pożary itd.) są małe i średnie firmy (np. Sevitel, EMAG, Hasso, Micon). W przypadku systemów dyspozytorskich i monitorowania pracy maszyn spotykamy już rozwiązania firm zarówno małych jak i dużych (np. Famur). Dostatecznie powszechną praktyką jest to, że producenci urządzeń wydobywczych traktują systemy monitorowania jako uzupełnienie dla dostarczanych przez nich urządzeń.

Duża liczba dostawców powoduje znaczące problemy z jakością, integracją i poprawną interpretacją gromadzonych danych. Oczywiście dostępne na rynku systemy dyspozytorskie (np. THOR [1.3]) pozwalają na wizualizację danych pochodzących z różnych systemów monitorowania, ale stopień integracji i dalszego wykorzystania danych jest niewielki. Gromadzone dane wykorzystywane są głównie do bieżącej (chwilowej) wizualizacji na planszach obrazujących określone miejsca kopalni. Często ten sam czujnik funkcjonuje w różnych systemach

pod różnymi nazwami, może się również zdarzyć, że w różnych systemach identyczny czujnik posiada różną konfigurację (np. różne wartości progów alarmowych). W praktyce, głębsza integracja systemów monitorowania następuje poprzez rozwiązania ad hoc, które dostosowane są do danej sytuacji. W przypadku zaistnienia sytuacji niepożądanych (np. poważnych awarii maszyn górniczych, niekontrolowanych i awaryjnych wyłączeń energii elektrycznej, pożarów, wypadków lub katastrof) dyspozytorzy lub odpowiednie komisje eksperckie podejmują próbę wyjaśnienia zaistniałej sytuacji m.in. poprzez analizę danych historycznych. Najczęściej dzieje się to poprzez eksport tych danych (często oddzielnie z każdego systemu pomiarowego) do arkusza kalkulacyjnego, gdzie następuje ich czasochłonna integracja, oczyszczanie i analiza.

Problemy związane z integracją i podnoszeniem jakości danych pojawiają się również w innych obszarach związanych z systemami monitorowania procesów przemysłowych. Przykładem może być stosunkowo niewielka integracja systemów monitorowania i zarządzania produkcją z systemami utrzymania ruchu (CMMS – Computerised Maintenance Management Systems). W Polsce zaledwie klika spośród kilkunastu systemów tej klasy jako integralny element wdrożenia oferuje integrację z systemami MES/SCADA, a informacje pochodzące z tych systemów wykorzystywane są do określania rzeczywistego stanu urządzeń (np. analiza liczby włączeń i wyłączeń) [1.10].

Można postawić tezę, że w interesującym nas obszarze zastosowań — systemów informatycznych wspomagających zadanie monitorowania procesów produkcyjnych i zagrożeń z nimi związanych — obserwuje się rosnące zainteresowanie (zarówno dostawców jak i odbiorców) metodami umożliwiającymi lepsze wykorzystanie gromadzonych danych [1.4, 1.5, 1.7–1.9, 1.11, 1.13, 1.14]. Firmy dostarczające informatyczne systemy monitorowania szukają swojej przewagi konkurencyjnej w wyposażeniu ich w metody inżynierii wiedzy, modelowania i analizy danych.

W dalszych rozdziałach niniejszej książki zaprezentowane zostaną wybrane elementy systemu wspomagania decyzji DISESOR. System jest efektem realizacji projektu badawczo-rozwojowego finansowanego przez Narodowe Centrum Badan i Rozwoju w ramach Programu Badań Stosowanych. Celem projektu było opracowanie metod i narzędzi umożliwiających utworzenie szkieletowego systemu wspomagania decyzji. System, po wyposażeniu w wiedzę dziedzinową oraz na podstawie analizy zgromadzonych danych, wspomaga użytkowników systemów dyspozytorskich i systemów monitorowania (ozn. SDM) w podejmowaniu decyzji mających znaczenie dla prawidłowego przebiegu monitorowanych procesów. Głównym przeznaczeniem opracowanych metod jest wspomaganie i rozszerzenie funkcjonalności SDM stosowanych w przemyśle wydobywczym, a szczególnie w górnictwie węgla kamiennego. Nadrzędną ideą projektu było opracowanie metod i narzędzi, które rozszerzą funkcjonalność SDM o możliwości oceny stopnia zagrożenia (np. wypadkami, katastrofami) i identyfikacji ryzyka (np. uszkodzenia maszyny, przestojów technologicznych) związanego z realizacją procesów technologicznych.

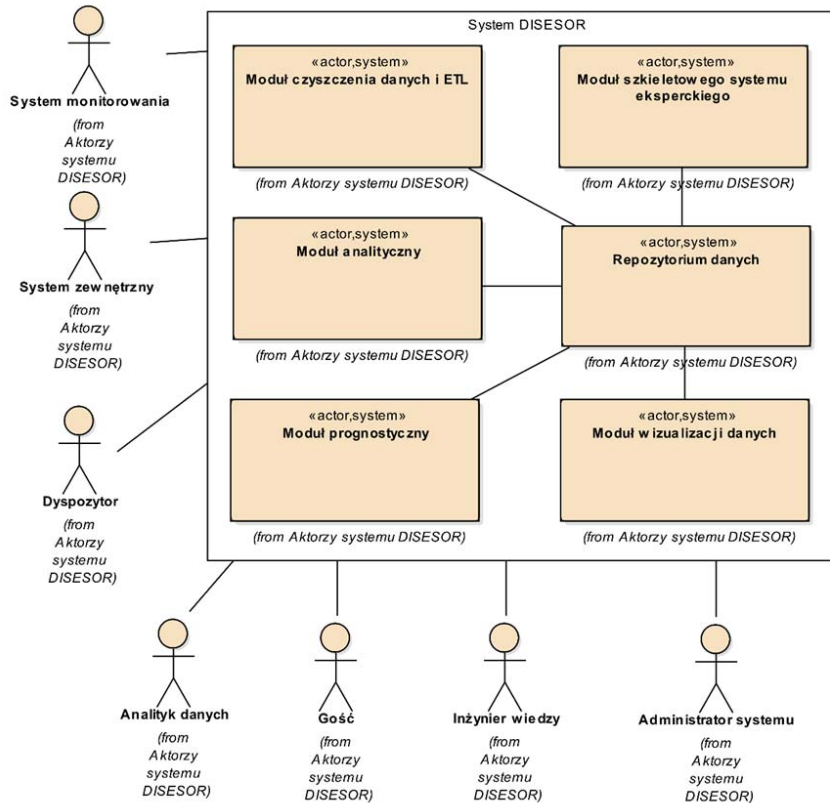
System jaki udało się zrealizować i jaki przedstawiony zostanie w dalszej części książki ma budowę szkieletową, dzięki temu może być zastosowany również w innych obszarach.

1.2 Wybrane aspekty architektury systemu DISESOR

Na rysunku 1.1 przedstawiono diagram kontekstowy pokazujący granice systemu DISESOR, źródła i odbiorców danych/informacji oraz jego główne wejścia i wyjścia. System ma strukturę modułową, w skład której wchodzi następujące komponenty: repozytorium danych, moduł czyszczenia, edycji i transformacji danych, moduł analityczny, moduł prognostyczny, moduł ekspercki oraz moduł wizualizacji danych. System współdziała z otoczeniem, którego głównymi aktorami są: administrator systemu, którego zadaniem jest dbanie o bezpieczeństwo wiedzy, informacji oraz danych przetwarzanych i gromadzonych przez system; analityk danych, który jest użytkownikiem odpowiedzialnym za przygotowanie modeli analitycznych i prognostycznych; inżynier wiedzy czyli osoba zajmująca się uzupełnianiem bazy wiedzy systemu poprzez tworzenie schematów wnioskowania; dyspozytor, czyli docelowy użytkownik systemu (end-user); systemy monitorowania, które są źródłem danych dla systemu DISESOR, a jednocześnie uzyskują z niego informację zwrotną jaką są np. stwierdzenia dotyczące wyników wnioskowania diagnostycznego lub wyniki oceny zagrożeń, itd.; gość, który jest użytkownikiem specjalnym np. takim, który ma dostęp do systemu w celach demonstracyjnych lub szkoleniowych; systemy zewnętrzne, które mogą być źródłem dodatkowych informacji/danych (np. kontekstowych) na temat działania urządzeń i maszyn oraz procesów technicznych i zagrożeń naturalnych. Przykładem w tym ostatnim przypadku mogą być serwisy informacyjne udostępniające dane pogodowe, które można brać pod uwagę podczas procesu wnioskowania na temat potencjalnych zagrożeń.

Niezwykle ważnym komponentem systemu jest repozytorium danych, które zaprojektowane zostało w taki sposób, aby możliwe było zasilenie go szerokim spectrum danych pomiarowych. Poprzez zdefiniowanie warstwy metadanych opisujących strukturę przedmiotu monitorowania możliwa jest jednoznaczna (niezależna od opisu w poszczególnych systemach pomiarowych) lokalizacja urządzeń i czujników, a przez to powiązanie mierzonych przez nie wielkości. Repozytorium jest również źródłem danych/informacji dla pozostałych komponentów systemu.

Za pomocą modułu analitycznego realizowane są zadania analizy zgromadzonych danych historycznych (off-line). Głównym użytkownikiem modułu jest analityk danych. W ramach modułu udostępniono wiele dedykowanych operatorów analitycznych, umożliwiających w sposób szybki i stosunkowo nieskomplikowany realizować zadania tworzenia modeli prognostycznych. Wyniki generowane przez ten moduł zasilają repozytorium danych dopiero po zaakceptowaniu przez użytkownika. Moduł ten wspiera zatem użytkownika w decydowaniu o tym co jeszcze warto monitorować i prognozować oraz dostarcza dodatkowych informacji, które mogą zostać wykorzystane dla wzbogacenia wiedzy systemu ekspertowego lub do analiz porównawczych. Dzięki modułowi analitycznemu możliwe jest: uzyskanie



Rys. 1.1. Diagram kontekstowy systemu DISESOR

informacji o tym czy wartości pomiarowe (oryginalne lub zagregowane) związane z danym aspektem monitorowanego procesu zmierzają w pożądanym lub niepożądanym kierunku, czy wartości te są powiązane z innymi wartościami oraz czy w świetle zgromadzonych danych można zidentyfikować przyczynę zarejestrowanych zmian. W module zdefiniowane są operatory pozwalające w sposób półautomatyczny realizować pewne podstawowe zadania (np. „prognozuj zagrożenie metanowe”, „prognozuj zagrożenia sejsmiczne”, „zbuduj model prognostyczny”, „śledź trend”). Moduł zawiera również dużą liczbę narzędzi umożliwiających realizację zaawansowanych, wymagających bardzo specjalistycznej wiedzy, procesów analitycznych. Użytkownik inicjować może zadania analityczne poprzez wybór i parametryzację predefiniowanych operacji (m.in. wybór zmiennych podlegających porównaniu, sposób agregacji dostępnych danych pomiarowych, określenie progów atrakcyjności odkrytych zależności, częstotliwości porównań itd.). Przedstawiona funkcjonalność pozwala na identyfikację zmian zachodzących w monitorowanych procesach i urządzeniach w postaci jawnej, możliwej do inter-

pretacji przez użytkownika. Podstawową reprezentację odkrytych zależności stanowią reguły, fakty oraz proste wskaźniki statystyczne ilustrujące rodzaj i siłę zidentyfikowanych zależności.

Kolejnym, ważnym elementem systemu jest moduł prognozowania, którego zadaniem jest wdrażanie do praktycznego działania modeli utworzonych w module analitycznym. Każdy model prognostyczny odpowiedzialny jest za prognozowanie wskazania konkretnego czujnika, rzeczywistego lub wirtualnego. Wirtualny czujnik to model, który służy np. do prognozowania stanu technicznego wybranej maszyny lub urządzenia. W module prognostycznym dokonuje się konfiguracji parametrów związanych z monitorowaniem jakości prognoz oraz określa się zasady na których — w przypadku spadku jakości prognoz poniżej pewnego minimum — model prognostyczny zostaje zaktualizowany. Zadaniem modułu prognostycznego jest również zarządzanie aktualnymi i nieaktualnymi modelami prognostycznymi. Do tego celu stosowane jest wewnętrzne repozytorium danych środowiska RapidMiner. Niezwykle ważną rolą modułu prognostycznego jest również przyjmowanie danych napływających bezpośrednio z systemów monitorowania, przetwarzanie ich zgodnie z wymaganiami danego modelu prognostycznego i obliczanie dla nich aktualnej wartości prognozy. Na rysunku 1.2 widać wyraźnie, że moduł prognostyczny przyjmuje również dane bezpośrednio z systemów monitorowania. Są to dane surowe, które nie mogą być przekazane bezpośrednio do modelu prognostycznego, gdyż ten zazwyczaj operuje na danych przetworzonych. Analityk opracowując model prognostyczny dedykowany dla danego czujnika dokonuje szeregu przekształceń danych (agregacji, wygładzania, ekstrakcji nowych i selekcji najbardziej istotnych cech, etc.). Wynikowy model prognostyczny przekazywany do modułu prognostycznego posiada zapisane wszystkie przekształcenia jakie wykonano na danych surowych. Moduł prognostyczny przetwarza dane surowe zgodnie z zapisami adekwatnymi dla każdego konkretnego modelu. Dzięki takiemu rozwiązaniu modele prognostyczne utworzone w module analitycznym można stosować do danych napływających on-line. Przekazanie danych napływających on-line bezpośrednio do modułu prognostycznego (z pominięciem repozytorium danych) przyspiesza w znaczącym stopniu czas podejmowania decyzji.

W systemie przewidziano również moduł szkieletowego systemu ekspertowego, którego przeznaczeniem jest diagnozowanie obiektów technicznych (urządzeń i maszyn) w trybie on-line lub off-line, nadzorowanie prawidłowego przebiegu procesów oraz wspomaganie podejmowania decyzji przez dyspozytora/eksperta (zarówno w odniesieniu do stanu technicznego rozpatrywanych obiektów jak i sytuacji kryzysowych wynikających ze zmierzania procesu w niepożądanym kierunku). Zaprojektowany i zrealizowany szkieletowy system ekspertowy umożliwia intuicyjne pozyskanie i zapisanie wiedzy w różnej formie oraz przeprowadzenie procesu wnioskowania z wykorzystaniem różnych metod elementarnych (wnioskowanie klasyczne w oparciu o reguły ostre i fakty, wnioskowanie probabilistyczne z zastosowaniem sieci Bayesa, czy też wnioskowanie posybilistyczne w oparciu o reguły rozmyte). Możliwość zapisu wiedzy w różnej postaci oraz zastosowany sposób wnioskowania (uwzględniający różne sposoby reprezentacji

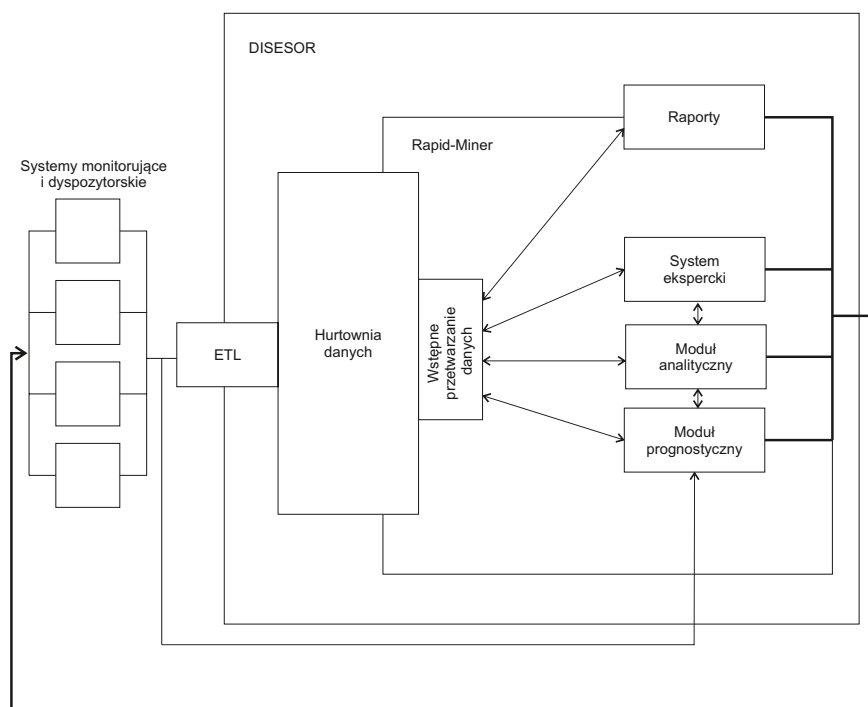
wiedzy) umożliwia wielowątkowe i wieloaspektowe wnioskowanie o stanie technicznym urządzeń, maszyn i procesów.

Ostatnim elementem systemu jest moduł wizualizacji i raportowania, którego zadaniem jest zaprezentowanie użytkownikowi w sposób czytelny gromadzonych danych pomiarowych oraz informacji pochodzących z pozostałych składowych systemu. Wizualizacja odbywa się za pośrednictwem systemu zewnętrznego jakim jest system dyspozytorski THOR. Dzięki takiemu rozwiązaniu użytkownik ma do dyspozycji narzędzia do edycji plansz oraz map wizualizacyjnych. Tak naprawdę wizualizacja nie jest jednak głównym zadaniem systemu DISESOR. W zasadzie każdy system współpracujący z DISESOREm, posiadający możliwości wizualizacji, może prezentować sygnały generowane przez system. Z punktu widzenia systemów monitorowania i systemów dyspozytorskich sygnałami takimi są: prognozowane wartości (wskazania) czujników, stan techniczny (aktualny bądź prognozowany) maszyn i urządzeń. Sygnały generowane przez system DISESOR mogą być traktowane jako tzw. wirtualne czujniki, które systemy monitorowania widzą jako kolejne źródło danych. W warstwie prezentacji wyników analizy DISESOR posiada własne plansze ilustrujące np. trafność generowanych prognoz, graficzną reprezentację gromadzonych danych (wykresy), itd. Z kolei moduł raportowania pozwala na przygotowanie szeregu raportów podsumowujących dane i informacje zgromadzone w repozytorium. Uzyskano w ten sposób funkcjonalność łączącą rozwiązania SCADA, MES, GIS/FM [1.2] z narzędziami Business Intelligence.

1.3 Implementacja systemu

Analizując różne możliwości praktycznej realizacji projektu, a więc możliwość jego implementacji najbardziej zbliżonej do wersji przeznaczonej do komercjalizacji, rozważano kilka możliwych scenariuszy, od całkowitej implementacji każdego składnika systemu do wykorzystania gotowego narzędzia powielarnego, umożliwiającego dostosowanie go do potrzeb projektu. Zdecydowano się na rozwiązanie pośrednie, a więc wybór środowiska istniejącego, ale otwartego, umożliwiającego swobodne modyfikacje i programowe rozszerzanie jego funkcjonalności. Jako jądro systemu wybrano oprogramowanie RapidMiner [1.6], w wersji działającej na zasadzie open source i takim modelu licencjonowania, który w zasadzie bez dodatkowych kosztów związanych z zakupem środowiska pozwala na przyszłą komercjalizację naszego rozwiązania. Wybór środowiska miał wpływ na narzędzia programowe wykorzystane do realizacji projektu, a także na architekturę systemu. Środowisko RapidMiner zapewnia wiele predefiniowanych funkcji przetwarzania danych oraz ich analizy. Pozwala to zaoszczędzić sporo czasu na implementację podstawowej funkcjonalności projektowanego systemu. RapidMiner zapewnia również podstawowy graficzny interfejs użytkownika. Z drugiej strony RapidMiner jest całkowicie konfigurowalny, można także dowolnie zmieniać jego funkcjonalność (język Java). W corocznym raporcie Gartnera (Magic Quadrant for Advanced Analytics Platforms) RapidMiner znalazł się obok SAS i IBM wśród liderów zaawansowanych platform analitycznych. Wyko-

rzystanie RapidMinera umożliwia również korzystanie z rozwiązań zawartych w coraz popularniejszym w zastosowaniach komercyjnych pakiecie R do obliczeń statystycznych.



Rys. 1.2. Umieszczenie środowiska RapidMiner w strukturze systemu DISESOR.

RapidMiner jest aplikacją przeznaczoną w głównej mierze do realizacji zadań związanych z eksploracją danych. Środowisko to zostało wyposażone w przejrzysty interfejs graficzny pozwalający na programowanie z wykorzystaniem metody przeciągnij i upuść. Dla bardziej zaawansowanych użytkowników systemu możliwe jest również skorzystanie z bibliotek programistycznych w sposób klasyczny. Dostępne wtyczki/operators, obszerna dokumentacja oraz otwarty kod źródłowy pozwala na rozbudowę funkcjonalności aplikacji w bardzo szerokim zakresie. Implementację kluczowych modułów systemu DISESOR zrealizowano w postaci nowych wtyczek aplikacji RapidMiner. W ten sposób uzyskano oprogramowanie będące szkieletowym systemem wspomagania decyzji wyposażonym w gotowe narzędzia umożliwiające pozyskanie wiedzy niezbędnej do poprawnego działania tego systemu. Dla celów realizacji systemu DISESOR opracowano również polską wersję środowiska.

Na rysunku 1.2 przedstawiono architekturę systemu DISESOR pokazującą umiejscowienie środowiska RapidMiner w strukturze systemu DISESOR.

Do realizacji repozytorium danych wykorzystano system zarządzania bazą danych PostgreSQL (w wersji umożliwiającej równoległą realizację zapytań). Przed podjęciem decyzji o takim właśnie wyborze przeanalizowano kilka rozwiązań dedykowanych dla przetwarzania danych strumieniowych (m.in. rozwiązania z obszaru Time Series Databases) i kolumnowych baz danych (np. Infobright). Testy wydajnościowe polegające na pomiarze czasu odpowiedzi na zapytania testowe pokazały, że odpowiednio skonfigurowany serwer PostgreSQL jest rozwiązaniem bardziej uniwersalnym o porównywalnej wydajności przetwarzania dużych zbiorów danych.

Do realizacji procesów ETL i raportowania zostały wybrano również otwarte rozwiązania Talend Open Studio [1.1] i BIRT [1.12].

Literatura

- [1.1] R. Barton. *Talend Open Studio Cookbook*. Packt Publishing Ltd, 2013.
- [1.2] S. Boyer. *Scada: Supervisory Control And Data Acquisition*. ISA: The Instrumentation, Systems, and Automation Society, 2009.
- [1.3] A. Dylong. System dyspozytorski THOR — nowa jakość w zakresie wizualizacji pracy i bezpieczeństwa zakładu górniczego. *Szkola Eksploatacji Podziemnej*, s. 67–74, 2013.
- [1.4] T. V. C. E.V. Chekushina, A.E. Vorobev. C. wang, z. wang. *Journal of Software*, 10(5):1114–1120, 2010.
- [1.5] T. V. C. E.V. Chekushina, A.E. Vorobev. Use of expert systems in the mining. *Middle-East Journal of Scientific Research*, 1(18):1–3, 2013.
- [1.6] M. Hofmann, R. Klinkenberg. *RapidMiner: Data Mining Use Cases and Business Analytics Applications*. Chapman and Hall/CRC, 2013.
- [1.7] J. Kabiesz. Effect of the form of data on the quality of mine tremors hazard forecasting using neural networks. *Geotechnical & Geological Engineering*, 24(5):1131–1147, 2006.
- [1.8] J. Korbicz, J. M. Koscielny, Z. Kowalczyk, W. Cholewa. *Fault diagnosis: models, artificial intelligence, applications*. Springer Science & Business Media, 2012.
- [1.9] M. Kozielski, M. Sikora, Ł. Wróbel. Disesor-decision support system for mining industry. *Computer Science and Information Systems (FedCSIS), 2015 Federated Conference on*, s. 67–74. IEEE, 2015.
- [1.10] A. Loska. Wybrane aspekty komputerowego wspomaganie zarządzania eksploatacją i utrzymaniem ruchu. *Monografia. Oficyna Wydawnicza Polskiego Towarzystwa Zarządzania Produkcją, Opole*, 2012.
- [1.11] W. Moczulski, K. Cyran, M. Januszka, P. Novak, A. Timofiejczuk. Telerescuer — an innovative robotized system for supporting mining rescuers by inspecting roadways affected by catastrophes. *24th World Mining Congress*, 2016.
- [1.12] D. Peh, N. Hague, J. Tatchell. *BIRT: A Field Guide (3rd Edition)*. Addison–Wesley Professional, 2011.

-
- [1.13] P. Stefaniak, R. Zimroz, W. Bartelmus, M. Hardygóra. Computerised decision-making support system based on data fusion for machinery systems management and maintenance. *Applied Mechanics and Materials*, 683:108–113, 2014.
- [1.14] R. Zimroz, M. Hardygóra, R. Błazej. Maintenance of belt conveyor systems in Poland - an overview. *Proceedings of the 12th International Symposium Continuous Surface Mining - Aachen*, s. 21–30. Springer Verlag, 2015.