

# ROLA I POTENCJAŁ DUŻYCH ZBIORÓW DANYCH W UTRZYMANIU RUCHU

**Artur Blaszczyk**

Politechnika Łódzka

Wydział Zarządzania i Inżynierii Produkcji

*artur.blaszczyk@p.lodz.pl*

**Streszczenie:** *Aplikacje gromadzące dane zyskują coraz większe znaczenie dla procesów produkcyjnych i utrzymania ruchu. Podczas codziennych operacji generowane są duże liczby plików i ich obfitość często paradoksalnie utrudnia zdolność do wyciągania użytecznej wiedzy. Ponadto duża liczba danych przechowywanych w często niepołączonych ze sobą bazach utrudnia ręczne analizowanie cennych wiadomości dla podejmowania decyzji. Współcześnie potrzebne są nowe narzędzia i techniki, które mogą w sposób inteligentny analizować zebrane informacje i dostarczać użytecznej wiedzy na potrzeby utrzymania ruchu. W dokumencie tym zbadano rolę danych dla funkcji konserwacji i sytuacje, w których nieprawidłowe zarządzanie nimi może powodować problemy zamiast dostarczać rozwiązania.*

**Słowa kluczowe:** utrzymanie ruchu, CMMS, zarządzanie zasobami.

## 1. Wprowadzenie

Dane są ważnym narzędziem w rękach personelu utrzymania ruchu, a proces ich pozyskiwania stał się prostszy i ekonomicznie przystępniejszy wraz z postępem technologicznym w systemach sprzętowych i oprogramowaniu. Duża ilość plików generowanych i gromadzonych przez służby zarządzające aktywami często nie jest wykorzystywana na pożądanym poziomie. W literaturze można napotkać dowody świadczące o tym, że większość organizacji dysponuje znacznie większymi zbiorami danych niż są w stanie wykorzystywać, a jednocześnie nie mają danych, których potrzebują [1]. Ponadto występują problemy związane z nieprawidłową jakością plików, co prowadzi do ich dezaktualizacji i stają się przestarzałe w organizacji. Pomimo tego, że rosnąca liczba wbudowanych systemów doprowadziła do wybuchu ery generowania danych, to wydaje się, że na szczeblu kierowniczym ludzie nie są przekonani, że posiadają wystarczająco dużo prawidłowych, wiarygodnych, spójnych i aktualnych informacji, na podstawie których mogliby podejmować decyzje [2]. Firmy z każdej branży zmagają się z trudnościami, a czasem wręcz bankrutują – w wyniku słabego dostępu do

danych i zarządzania nimi, niemożności ich przetłumaczenia na cenne informacje oraz ich złej jakości [3].

Współcześnie istnieje kilka poważnych trudności napotykanych w przypadku utrzymania ruchu bazującego na stanie technicznym urządzeń (CBM – ang. *Condition Based Maintenance*). Ze względu na ogólnie dużą ilość generowanych danych, może zaistnieć konieczność ich gromadzenia z zasobów rozproszonych na znacznym obszarze i w konsekwencji pojawia się problem ich akwizycji. Ponadto istnieje możliwość potrzeby scalenia plików celem dostarczenia wszelkich użytecznych informacji, z uwzględnieniem czasu potrzebnego na ich pozyskanie z dodatkowych źródeł, zintegrowania z resztą w celu bardziej zrozumiałej interpretacji oraz – na koniec – dostępności eksperta, który przekształci je na użyteczne informacje dla utrzymania ruchu. Dostęp do wykwalifikowanych specjalistów jest niestety niewielki, a zatem nawet w przypadku uruchomienia programu monitorowania stanu, awarie nadal często występują, nie realizując samego celu, dla którego dokonuje się inwestycji w CBM [4][5].

W celu rozwiązania powyższych problemów w obszarze zainteresowań znalazły się różne typy technologii informatycznych. Sztuczna inteligencja (AI – ang. *Artificial Intelligence*) dla podejmowania decyzji w utrzymaniu ruchu zaczęła pojawiać się w latach 80. w formie systemów eksperckich. W latach 90. natomiast obszarem zainteresowań stały się sztuczne sieci neuronowe i logika rozmyta. W literaturze dostępne są opracowania na temat sztucznej inteligencji [6][7] oraz jej zastosowania do monitorowania stanu [8][9]. Rozproszona AI była również stosowana do monitorowania stanu urządzeń po upowszechnieniu Internetu pod koniec lat 90. [10][11][12][13]. Od niedawna naukowcy zaczęli stosować technologie sieci i agentów w utrzymaniu ruchu, a przegląd tej kwestii został wydany przez Campos i Prakash [14], a następnie rozszerzony i zaktualizowany przez Campos [15]. Technologie te uzyskały szerszą akceptację ze względu na zdolność operatorów do działania na rozproszonych, otwartych środowiskach, takich jak Internet lub Intranet korporacyjny, a także dostęp do heterogenicznych i geograficznie oddalonych baz danych i źródeł informacji [16][17].

Obecnie, przy zwiększonej pojemności dysków przedsiębiorstwa gromadzą ogromną ilość plików. Odpowiedni wgląd w produkowane materiały stanowi wyzwanie oraz tworzy przewagę konkurencyjną dla firm, które są w stanie wykorzystać je optymalnie dla swoich korzyści [18]. Analiza dużych zbiorów danych w utrzymaniu ruchu znajduje się w fazie wprowadzania, gdzie przykładowe ramy i wytyczne są sugerowane oraz testowane w różnych sektorach przemysłu [19][20][21][22]. Implementacja powyższych systemów niesie ze sobą wiele możliwości, a także wyzwań ze względu na złożoność plików, co zostało podkreślone w niniejszej pracy.

## 2. Problematyka danych w zarządzaniu zasobami

Złożoność danych w zarządzaniu zasobami zlokalizowanymi w jednym miejscu lub rozproszonymi stanowi problem dla menedżerów i zaczyna się często z powodu niewłaściwego wyboru urządzeń i parametrów pomiarowych. W przedsiębiorstwach przemysłowych istnieje tendencja, aby mierzyć to, co łatwo mieralne, a nie to, co jest wymagane [23], co powoduje zebranie bezużytecznym danych, które nie mają wartości analitycznej. Te nieprawidłowe pliki należy rozumieć w koncepcji biznesowej jako marnotrawstwo oraz bezproduktywną pracę [3]. Głównym problemem jest ich przekształcenie w informacje, które mogą być wykorzystane do podejmowania znaczących decyzji w zakresie zarządzania.

Ponadto organizacje borykają się z problemami związanymi z kompatybilnością z powodu występowania oddzielnych pakietów oprogramowania komputerowego do zarządzania firmą i do zarządzania zasobami technicznymi. Systemy te najczęściej nie są w stanie się ze sobą porozumiewać, co doprowadziło do wprowadzenia pojęcia „wysp informacyjnych”, i sprawia problemy w dostarczaniu informacji w czasie rzeczywistym z systemów utrzymania ruchu do systemów biznesowych [2].

Czynnik ludzki w pozyskiwaniu danych ma zarówno pozytywne, jak i negatywne aspekty. Z jednej strony człowiek ma większą tolerancję na niejednoznaczność i może przyswajać heterogeniczność, a bogactwo naturalnego języka w rzeczywistości niejednokrotnie dostarcza cennej głębi. Jednakże algorytmy analityczne maszyn oczekują jednorodnych danych i nie radzą sobie w odczytywaniu niuansów [24]. Może to prowadzić do problemów, w których operatorzy, którzy nie przechodzą odpowiednich szkoleń, odczytują niepoprawne i nieistotne dane, co prowadzi do tworzenia błędnych wniosków. Z drugiej strony część literatury podkreśla przewagę interwencji człowieka nad prostym bazowaniem na pomiarach z czujników. Czynniki ludzkie związane z gromadzeniem danych są ponadto bardziej wiarygodne ze względu na ścisłe związanie ze wskaźnikami własności i odpowiedzialności. Technicy i operatorzy zbierają próbki tylko wtedy, gdy uważają to za zasadne, a wyniki są udostępniane do konsultacji i wykorzystania [25]. Pomimo zdecydowanego wzrostu liczby czujników, które potrafią odczytywać i dostarczać dane, co pociąga za sobą dodatkowy koszt dla przedsiębiorstwa, wiele zakładów nadal rejestruje i aktualizuje je w systemach informatycznych w dużej mierze manualnie lub tylko półautomatycznie. Rozwiązanie technologiczne samo w sobie nie może poprawić integralności plików, ponieważ często wymaga systematycznych zmian instytucjonalnych [26].

Powody złożoności danych wynikające z ich złej jakości można podsumować jako [2]:

- nieodpowiednie struktury zarządzające do zapewnienia pełnego, terminowego i dokładnego raportowania,
- niewystarczające szkolenia i procedury dla osób zaangażowanych w gromadzenie plików,
- rozdrobnienie i niespokojność wśród jednostek zbierających dane,
- wymóg dotyczący nowych metod zarządzania, które wykorzystują dokładne i istotne dane, aby wspierać dynamiczne środowisko zarządzające.

### 3. Obszar dużych zbiorów danych

Wielkość dużych zbiorów danych najłatwiej zrozumieć empirycznie. Przykładowo, każdorazowe przeprowadzenie eksperymentu przez Wielki Zderzacz Hadronów w CERN jest w stanie wygenerować 40 terabajtów (TB) danych, a 30 minut pracy silnika samolotu Boeing tworzy 10 TB informacji dotyczących operacji [27]. Ponadto każda pojedyncza podróż przez Ocean Atlantycki może przynieść nawet 640 TB plików, a takich odbywa się dziennie ponad 25 tysięcy [27]. Przeładowanie ilością danych staje się w dzisiejszych czasach problemem. Na poziomie organizacyjnym generowane są ogromne ich ilości, w tym strukturalne i niestrukturalne, trwałe i czasowe itd. [2]. Projekty dotyczące dużych zbiorów danych nieustannie się rozwijają ze względu na konieczność obsłużenia użytkowników na większą skalę i ich rosnące wymagania [28]. Dane mogą być filtrowane i kompresowane w zależności od wielkości i znaczenia bez ograniczenia możliwości prowadzenia analiz. Jednym z wyzwań jest stałe definiowanie tych filtrów w taki sposób, aby nie odrzucały użytecznych informacji z tego względu, że dane pierwotne są często zbyt obszerne i niemożliwe jest nawet przechowanie ich wszystkich [24]. Duże zbiory danych składają z obszernych, heterogenicznych i autonomicznych źródeł, które są poddawane analizie i poszukiwaniu relacji między nimi.

Powyższe cechy stanowią wyzwanie dla czerpania użytecznej wiedzy z dużych zbiorów danych. Wielowymiarowość i rozmiar próbek w plikach rodzi następujące problemy [29]:

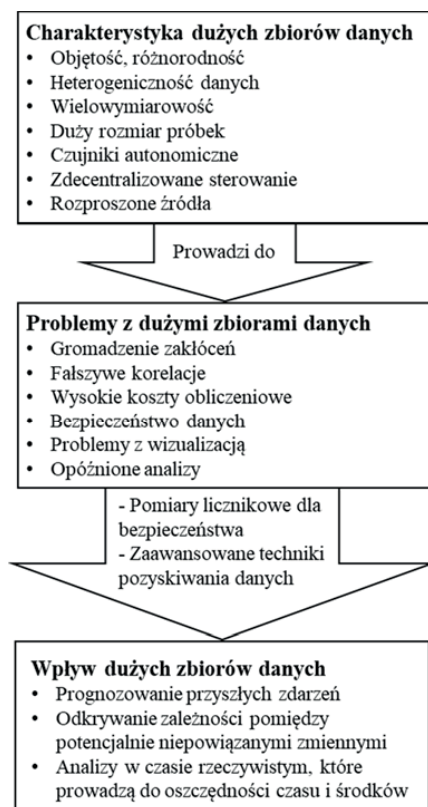
- nagromadzenie szumów i pozorne korelacje wynikające z wysokiej rozdzielczości,
- wysoki koszt obliczeniowy i niestabilność algorytmów ze względu na wysoką wymiarowość w dużych próbkach,
- zbieranie danych z różnych źródeł, w różnym czasie i przy użyciu różnych technologii, co tworzy problemy heterogeniczności i błędy statystyczne oraz wymaga opracowania bardziej adaptacyjnych i solidnych algorytmów.

Spojrzenie człowieka na duże zbiory danych również ulegało zmianie, gdy zaczęły się one stawać coraz obszerniejsze. W celu zmniejszenia luki pomiędzy

osobą a systemem tworzy się przyjazne dla użytkownika wizualizacje, które ukazują wyniki analiz w sposób intuicyjny, tak, aby możliwe było skuteczne zidentyfikowanie interesujących informacji. Jednym z wymagań jest odpowiednia wydajność systemu w czasie rzeczywistym, która pozwoli na szybkie przekształcenie zapytania użytkownika w odpowiednio przygotowaną odpowiedź [28][24].

#### 4. Korzyści wynikające z dużych zbiorów danych

Duże zbiory danych rewolucjonizują obecnie wiele dziedzin i znalazły zastosowanie, między innymi, w medycynie [30], ekonomii i finansach [29] oraz prognozowaniu rynku akcji [31]. Mają one również duży potencjał w utrzymaniu ruchu. W poprzednim rozdziale omówiono wyzwania związane z analizą dużych zbiorów danych ze względu na ich wielowymiarowość i rozmiary próbek. Jednak wysoka dokładność danych pozwala także na wykorzystanie ich do opracowania skutecznych metod pozwalających na przewidzenie przyszłych zdarzeń



Rys. 1. Duże zbiory danych w utrzymaniu ruchu

*Źródło: opracowanie własne.*

oraz jednocześnie uzyskanie wglądu w związek pomiędzy cechami a generowaną reakcją. Gdy dane są już obecne, to jedynymi czynnikami ograniczającymi osiągnięcie wspomnianych celów jest ich dostępność do analiz oraz istnienie technik analitycznych [29]. Ważnym aspektem jest wyciągnięcie wniosków z danej analizy w odpowiednich ramach czasowych, w których są one nadal istotne. W przypadku utrzymania ruchu analiza parametrów stanu maszyny powinna prowadzić do uzyskania prognozy o nadchodzącej awarii zanim ona wystąpi, ponieważ po fakcie informacja ta stanie się bezwartościowa.

W celu sprostania oczekiwaniom dotyczącym analiz dużych zbiorów danych wydajność obliczeniowa powinna znacznie wzrosnąć. Ponadto duże dane, ze względu na ich kompresowanie, mogą prowadzić do nieścisłości statystycznych, co skutkuje niejednokrotnie fałszywymi korelacjami. Taka zmiana paradygmatu doprowadziła do znacznych postępów w opracowywaniu szybkich algorytmów, które są skalowalne do wielkich danych o dużej wymiarowości, co umożliwiła współpraca różnych dziedzin, takich jak statystyka, optymalizacja i matematyka stosowana [29]. Można wnioskować, że wraz z coraz większym zaawansowaniem prac, luka pomiędzy danymi a informacjami będzie wypełniana. Na rysunku 1 podsumowano problemy związane z dużymi zbiorami danych oraz podkreślono potrzebę posiadania środków bezpieczeństwa i technik pozyskiwania plików w celu wykorzystania ich pełnego potencjału.

## **5. Dane gromadzone w utrzymaniu ruchu**

Badania dotyczące gromadzenia danych w utrzymaniu ruchu przeprowadzono w przedsiębiorstwach produkcyjnych o różnej wielkości i różnym czasie działania na rynku oraz reprezentujących różne branże i znajdujących się w różnych lokalizacjach. Firmy podzielono na trzy grupy w zależności od statusu implementacji komputerowych systemów zarządzania utrzymaniem ruchu (CMMS). Kwestionariusz ankietowy został zrealizowany przy pomocy formularza internetowego i użyto w nim pytań otwartych i zamkniętych, z możliwością jednokrotnego bądź wielokrotnego wyboru oraz wykorzystano odpowiednie skale nominalne, przedziałowe. Respondentami ankiety właściwej byli menedżerowie działów utrzymania ruchu i menedżerowie działów technicznych.

Wyniki badań opracowano przy zastosowaniu metod opisowych i wnioskowania statystycznego. W opisie badanych przedsiębiorstw oraz ich podgrup ze względu na fakt: implementacji CMMS, planowania implementacji CMMS lub nieplanowania implementacji CMMS w przyszłości, obliczono wskaźniki struktury. Dla porównania częstości występowania różnych odmian cech w podgrupach, badając zależności pomiędzy cechami, zastosowano test niezależności  $\chi^2$ .

Respondentom zadano pytanie dotyczące informacji gromadzonych w Dziale Utrzymania Ruchu. Wyniki zostały zebrane poniżej.

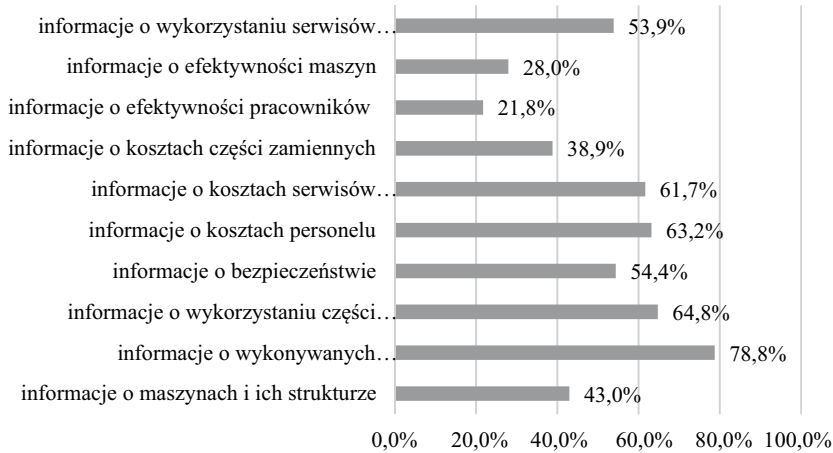
Analiza i porównanie kolejnych danych zbieranych w trzech grupach przedsiębiorstw wskazały istotną różnicę dla informacji o wykorzystaniu części zamiennych ( $p < 0,05$ ), informacji o bezpieczeństwie ( $p < 0,001$ ), informacji o kosztach personelu ( $p < 0,05$ ), informacji o kosztach serwisów zewnętrznych ( $p < 0,001$ ), informacji o efektywności pracowników ( $p < 0,05$ ) oraz informacji o wykorzystaniu serwisów zewnętrznych ( $p < 0,001$ ). Pozostałe informacje gromadzone przez Dział Utrzymania Ruchu nie różniły się istotnie w trzech porównywanych rodzajach firm ( $p > 0,05$ ).

Tabela 1. Informacje gromadzone w Dziale Utrzymania Ruchu

| Informacje gromadzone<br>w Dziale Utrzymania Ruchu        | CMMS                 |         |                                    |         |                                                |         | Razem<br>(n = 193) |       |
|-----------------------------------------------------------|----------------------|---------|------------------------------------|---------|------------------------------------------------|---------|--------------------|-------|
|                                                           | Wdrożony<br>(n = 84) |         | Wdrożenie<br>planowane<br>(n = 66) |         | Brak,<br>wdrożenie<br>nieplanowane<br>(n = 43) |         |                    |       |
|                                                           | n                    | Frakcja | n                                  | Frakcja | n                                              | Frakcja | n                  | %     |
| informacje o maszynach i ich<br>strukturze                | 40                   | 0,48    | 31                                 | 0,47    | 12                                             | 0,28    | 83                 | 43,0% |
| informacje o wykonywanych<br>działaniach (zlecenia pracy) | 84                   | 1,00    | 41                                 | 0,62    | 27                                             | 0,63    | 152                | 78,8% |
| informacje o wykorzystaniu<br>części zamiennych           | 74                   | 0,88    | 25                                 | 0,38    | 26                                             | 0,60    | 125                | 64,8% |
| informacje o bezpieczeństwie                              | 84                   | 1,00    | 7                                  | 0,11    | 14                                             | 0,33    | 105                | 54,4% |
| informacje o kosztach<br>personelu                        | 74                   | 0,88    | 21                                 | 0,32    | 27                                             | 0,63    | 122                | 63,2% |
| informacje o kosztach<br>serwisów zewnętrznych            | 84                   | 1,00    | 16                                 | 0,24    | 19                                             | 0,44    | 119                | 61,7% |
| informacje o kosztach części<br>zamiennych                | 35                   | 0,42    | 23                                 | 0,35    | 17                                             | 0,40    | 75                 | 38,9% |
| informacje o efektywności<br>pracowników                  | 25                   | 0,30    | 5                                  | 0,08    | 12                                             | 0,28    | 42                 | 21,8% |
| informacje o efektywności<br>maszyn                       | 30                   | 0,36    | 12                                 | 0,18    | 12                                             | 0,28    | 54                 | 28,0% |
| informacje o wykorzystaniu<br>serwisów zewnętrznych       | 79                   | 0,94    | 15                                 | 0,23    | 10                                             | 0,23    | 104                | 53,9% |
| inne                                                      | 0                    | 0,00    | 0                                  | 0,00    | 0                                              | 0,00    | 0                  | 0,0%  |

Źródło: opracowanie własne w oparciu o badania.

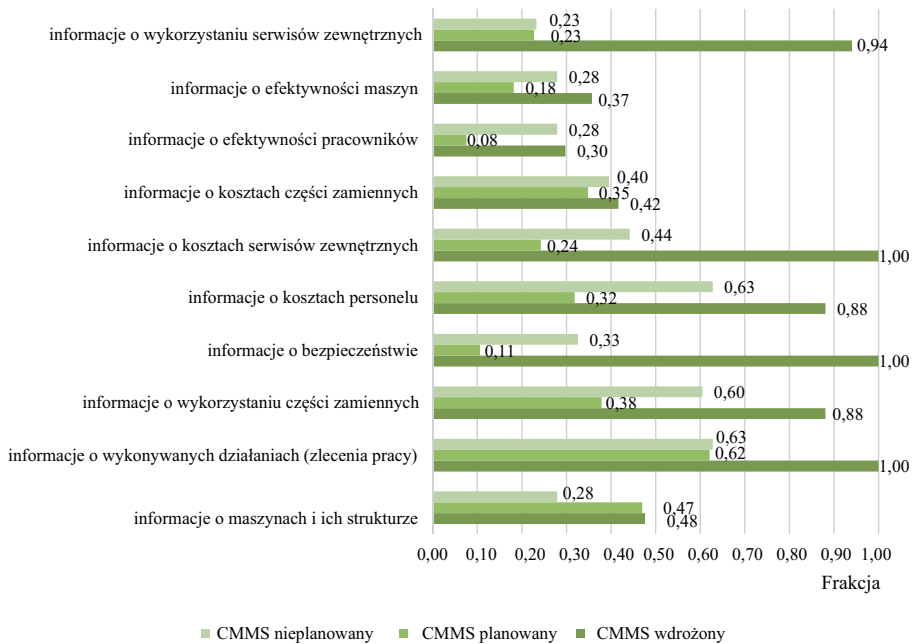
### Informacje gromadzone w Dziale Utrzymania Ruchu



Rys. 2. Informacje gromadzone w Dziale Utrzymania Ruchu

Źródło: opracowanie własne w oparciu o badania.

### Informacje gromadzone w Dziale Utrzymania Ruchu



Rys. 3. Informacje gromadzone w Dziale Utrzymania Ruchu

Źródło: opracowanie własne w oparciu o badania.



Na podstawie przeprowadzonych badań można stwierdzić, że gromadzenie informacji w Dziale Utrzymania Ruchu jest dodatnio skorelowane z implementacją CMMS. System ten gwarantuje odpowiednie ich uporządkowanie i zarządzanie nimi oraz umożliwia dostęp do danych wielu pracownikom jednocześnie. Przedsiębiorstwa, które pracują dłuższy czas z CMMS odnotowują znaczną optymalizację kosztów konserwacji, czasu pracy techników, a także usług świadczonych przez firmy podwykonawcze. Z rozmów z respondentami wynika, że jest to duża wartość dodana, która nie mogłaby zostać osiągnięta przy tradycyjnym zbieraniu danych w formie papierowej.

Ankietowani zostali zapytani również o występowanie w firmie systemu rejestracji czasu pracy i awarii środków trwałych. Otrzymane wyniki zostały zestawione w poniższej tabeli.

Tabela 2. Występowanie w firmie systemu rejestracji czasu pracy i awarii środków trwałych

| System rejestracji czasu pracy i awarii środków trwałych | CMMS     |         |                     |         |                              |         | Razem |         |
|----------------------------------------------------------|----------|---------|---------------------|---------|------------------------------|---------|-------|---------|
|                                                          | Wdrożony |         | Wdrożenie planowane |         | Brak, wdrożenie nieplanowane |         |       |         |
|                                                          | n        | Frakcja | n                   | Frakcja | n                            | Frakcja | n     | %       |
| Tak                                                      | 69       | 0,82    | 20                  | 0,30    | 5                            | 0,12    | 94    | 48,70%  |
| Nie                                                      | 15       | 0,18    | 46                  | 0,70    | 38                           | 0,88    | 99    | 51,30%  |
| Razem                                                    | 84       | 1,00    | 66                  | 1,00    | 43                           | 1,00    | 193   | 100,00% |

$$\chi^2 = 70,2; p < 0,001$$

Źródło: opracowanie własne w oparciu o badania.

Okazało się, że istnieje istotna różnica występowania w przedsiębiorstwach systemu rejestracji czasu pracy i awarii środków trwałych w zależności od faktu implementacji lub braku implementacji CMMS ( $p < 0,001$ ). Najczęściej system ten występuje w przedsiębiorstwach z wdrożonym informatycznym systemem zarządzania utrzymaniem ruchu (frakcja 0,82), następnie w organizacjach z planowanym wdrożeniem (frakcja 0,30), a najrzadziej w organizacjach, które wdrożenia CMMS nie planują (frakcja 0,12). W przypadku tych pierwszych wynika to z faktu oferowania niniejszej funkcjonalności przez system informatyczny.

Gromadzenie danych o awariach urządzeń jest szczególnie ważne we współczesnych przedsiębiorstwach ze względu na rotację pracowników. W przypadku wykorzystywania systemu informatycznego, osoby zatrudnione w dziale technicznym mają możliwość przejrzeć historyczne wpisy, które są nieocenionym źródłem wiedzy o maszynie i na ich podstawie zdiagnozować usterkę. Prawdopodobne jest, że przyczyny zostały już wcześniej określone i wystarczy je jedynie odszukać w programie.

Następnie zapytano respondentów o informacje dotyczące eksploatacji rejestrowane przez przedsiębiorstwo.

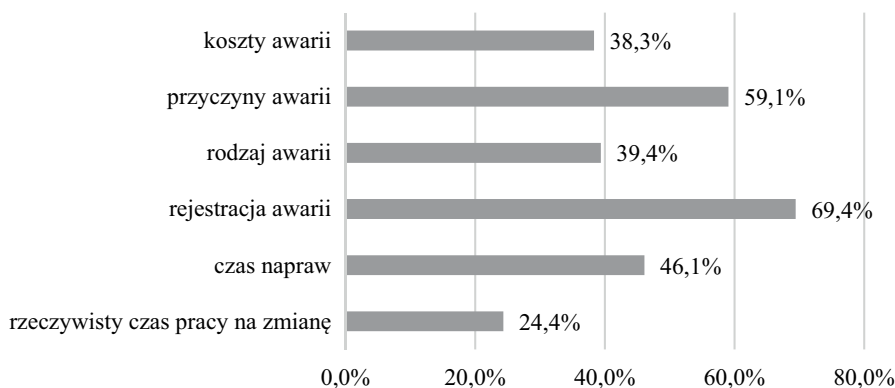
Tabela 3. Informacje dotyczące eksploatacji rejestrowane przez firmę

| Rejestracja eksploatacji         | CMMS                 |         |                                 |         |                                          |         | Razem<br>(n = 193) |       |
|----------------------------------|----------------------|---------|---------------------------------|---------|------------------------------------------|---------|--------------------|-------|
|                                  | Wdrożony<br>(n = 84) |         | Wdrożenie planowane<br>(n = 66) |         | Brak, wdrożenie nieplanowane<br>(n = 43) |         |                    |       |
|                                  | n                    | Frakcja | n                               | Frakcja | n                                        | Frakcja | n                  | %     |
| rzeczywisty czas pracy na zmianę | 25                   | 0,30    | 11                              | 0,17    | 11                                       | 0,26    | 47                 | 24,4% |
| czas napraw                      | 69                   | 0,82    | 11                              | 0,17    | 9                                        | 0,21    | 89                 | 46,1% |
| rejestracja awarii               | 84                   | 1,00    | 22                              | 0,33    | 28                                       | 0,65    | 134                | 69,4% |
| rodzaj awarii                    | 35                   | 0,42    | 27                              | 0,41    | 14                                       | 0,33    | 76                 | 39,4% |
| przyczyny awarii                 | 79                   | 0,94    | 13                              | 0,20    | 22                                       | 0,51    | 114                | 59,1% |
| koszty awarii                    | 35                   | 0,42    | 9                               | 0,14    | 30                                       | 0,70    | 74                 | 38,3% |
| inne                             | 0                    | 0,00    | 0                               | 0,00    | 0                                        | 0,00    | 0                  | 0,0%  |

Źródło: opracowanie własne w oparciu o badania.

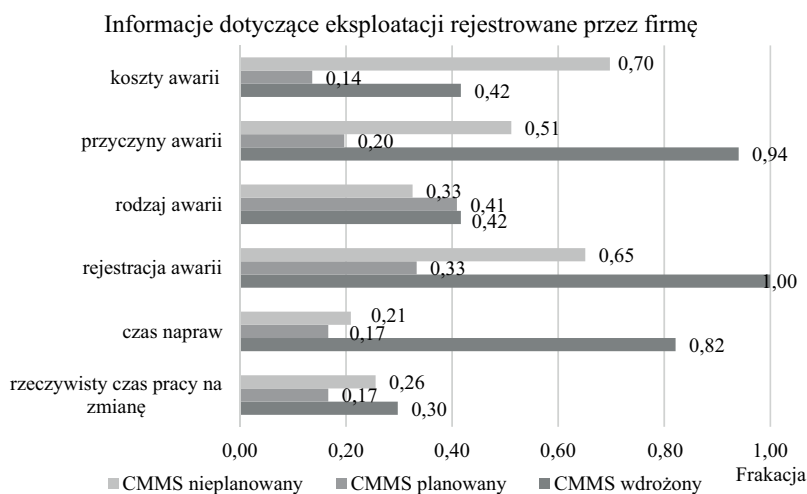
Jak wskazują wyniki, najczęściej zbierano informacje o wystąpieniu awarii (69,4%) oraz o przyczynach awarii (59,1%), a najrzadziej o rzeczywistym czasie pracy na zmianę (24,4%) i kosztach awarii (38,3%). Dwie ostatnie dane są istotne głównie ze względu na optymalizację budżetu działu utrzymania ruchu.

Informacje dotyczące eksploatacji rejestrowane przez firmę



Rys. 4. Informacje dotyczące eksploatacji rejestrowane przez firmę

Źródło: opracowanie własne w oparciu o badania.



Rys. 5. Informacje dotyczące eksploatacji rejestrowane przez firmę

Źródło: opracowanie własne w oparciu o badania.

Tabela 4. Porównanie częstości rejestracji poszczególnych informacji o eksploatacji

| Rejestracja eksploatacji         | Wartość testu chi <sup>2</sup> | Istotność p |
|----------------------------------|--------------------------------|-------------|
| rzeczywisty czas pracy na zmianę | 2,171                          | p>0,05      |
| czas napraw                      | 28,51                          | p<0,001     |
| rejestracja awarii               | 15,012                         | p<0,001     |
| rodzaj awarii                    | 0,49                           | p>0,05      |
| przyczyny awarii                 | 23,738                         | p<0,001     |
| koszty awarii                    | 15,941                         | p<0,001     |

Źródło: opracowanie własne w oparciu o badania.

Analiza i porównanie kolejnych rejestrowanych przez firmę danych o eksploatacji w trzech grupach przedsiębiorstw wskazały istotną różnicę dla czasu naprawy ( $p<0,001$ ), rejestracji awarii ( $p<0,001$ ), przyczyn awarii ( $p<0,001$ ) oraz kosztów awarii ( $p<0,001$ ). W przypadku informacji o rejestracji awarii, występowały zawsze tam, gdzie wdrożono CMMS (frakcja 1,00). Podobnie wyglądała sytuacja dla informacji o przyczynach awarii i czasie naprawy – frakcje odpowiednio 0,94 i 0,82. Dane o kosztach awarii najczęściej były rejestrowane w firmach nieplanujących implementacji CMMS (frakcja 0,70). Pozostałe rejestrowane w organizacjach informacje o eksploatacji miały zbliżoną częstość, która nie różniła się istotnie ( $p>0,05$ ).

Występująca powszechnie presja we współczesnych przedsiębiorstwach na redukcję przestojów produkcji oraz generowanie corocznych oszczędności niejednokrotnie zmusza działy utrzymania ruchu do poszukiwania nowych metod i technik. Nieocenionym źródłem informacji są w takim przypadku CMMS, które umożliwiają analizowanie zgromadzonych w nich danych i wyciąganie wniosków. Organizacje analizują m.in. czas pracy techników i koszty konserwacyjne, dążąc do ich zoptymalizowania.

## 6. Podsumowanie

W niniejszym rozdziale przestudiowano problematykę dużych zbiorów danych oraz korzyści z nich płynące. W celu utrzymania konkurencyjności przez przedsiębiorstwa ważne jest zrozumienie ich wartości i jak najlepsze prowadzenie analiz, prowadzących do uzyskania wartościowych informacji. Wnioski te są szczególnie ważne z punktu widzenia utrzymania ruchu, ponieważ pozwalają na zdefiniowanie optymalnych strategii konserwacyjnych, co przekłada się na obniżenie awaryjności parku maszynowego oraz kosztów prowadzonych prac.

Współczesne opracowania pokazują, że obecnie tworzy się i wykorzystuje różne technologie informacyjno-komunikacyjne (ICT), które umożliwiają przedsiębiorstwom organizowanie procesów. Należą do nich m.in. komputerowe systemy zarządzania utrzymaniem ruchu (CMMS). W przypadku ICT, takich jak pozyskiwanie danych i systemy dużych zbiorów danych, ważne jest zrozumienie celów i wymagań, aby było możliwe dobranie i użycie odpowiednich metod analitycznych. W związku z tym pracownicy utrzymania ruchu, którzy są ekspertami w dziedzinie wnioskowania, stają się kluczowymi osobami w organizacji, ponieważ informacje uzyskane w wyniku ich pracy mogą przyczynić się do zmniejszenia awaryjności parku maszynowego, a co za tym idzie – generowania oszczędności.

## Literatura

- [1] Levitan A.V., Redman T.C., *Data as a resource: properties, implications and prescriptions*, "Sloan Management Review" 1998, t. 40, pp. 89-101.
- [2] Koronios A., Lin S., Gao J., (2005), *A Data Quality Model for Asset Management in Engineering Organizations*, Proceedings of the 10th International Conference on Information Quality, Massachusetts Institute of Technology, Cambridge .
- [3] Gilliland G. et al., (2011), *Creating Value through data integrity: A pragmatic approach*, BCG Perspectives.
- [4] Rao J.S., Zubair M., Rao C., *Condition monitoring of power plants through the Internet*, "Integrated Manufacturing Systems" 2003, t. 14, pp. 508-517.

- [5] Prakash O., (2006), *Asset management through condition monitoring – how it may go wrong: a case study*, Proceedings of the 1st World Congress on Engineering Asset Management.
- [6] Kalogirou S.A., *Artificial intelligence for the modeling and control of combustion processes: A review*, “Progress in Energy and Combustion Science” 2003, t. 29, pp. 515-566.
- [7] Liao S.H., *Expert system methodologies and applications – a decade review from 1995 to 2004*, “Expert Systems with Applications” 2005, t. 28, pp. 93-103.
- [8] Warwick K., Ekwue A.O., Aggarwal R., (1997), *Artificial Intelligence Techniques in Power Systems*, Institution of Electrical Engineers, Stevenage.
- [9] Wang K., *Intelligent Condition Monitoring and Diagnosis System A Computational Intelligent Approach*, “Frontiers in Artificial Intelligence and Applications” 2003, t. 93.
- [10] Rao M., Yang H., Yang H., (1996), *Integrated distributed intelligent system for incident reporting in DMI pulp mill, Success and Failures of Knowledge-Based Systems in Real-World Applications*, Proceedings of the First International Conference, pp. 169-178.
- [11] Rao M., Zhou J., Yang H., (1998) *Architecture of integrated distributed intelligent multimedia system for on-line real-time process monitoring*, SMC’98 Conference Proceedings, 1998 IEEE International Conference on Systems, Man, and Cybernetics, pp. 1411-1416.
- [12] Rao M., Zhou J., Yang H., *Integrated distributed intelligent system architecture for incidents monitoring and diagnosis*, “Computers in Industry” 1998, t. 37, pp. 143-148.
- [13] Reichard K. M., Van Dyke M., Maynard K., (2000), *Application of sensor fusion and signal classification techniques in a distributed machinery condition monitoring system*, Proceedings of SPIE – The International Society for Optical Engineering, t. 4051, pp. 329-336.
- [14] Campos J., Prakash O., *Information and communication technologies in condition monitoring and maintenance*, “Information Control Problems in Manufacturing” 2006, pp. 3-8.
- [15] Campos J., *Survey Paper: Development in the Application of ICT in Condition Monitoring and Maintenance*, “Computers in Industry” 2009, t. 60, pp. 1-20.
- [16] Sycara K.P., *MultiAgent Systems*, “AI Magazine” 1998, t. 19.
- [17] Feng J.Q. et al., (2002), *A multi-agent based intelligent monitoring system for power transformers in distributed substations*, International Conference on Power System Technology Proceedings., t. 3, pp. 1962-1965.
- [18] Assuncao M. D. et al., *Big Data computing and clouds: Trends and future directions*, “Journal of Parallel and Distributed Computing” 2015, pp. 3-15.
- [19] Xu B., Kumar S., (2015), *Big Data Analytics Framework for System Health Monitoring*, IEEE International Congress on Big Data.
- [20] Fumeo E., Oneto L., Anguita D., *Condition Based Maintenance in Railway Transportation Systems Based on Big Data Streaming Analysis*, “Procedia Computer Science” 2015, t. 53, pp. 437-446.

- [21] Mohamed A., Hamdi M. S., Tahar S., (2015), *A machine learning approach for big data in oil and gas pipelines*, 2015 International Conference on Future Internet of Things and Cloud (FiCloud), 2015 International Conference on Open and Big Data (OBD). Proceedings, IEEE Computer Society.
- [22] Nunez A. et al., (2014), *Facilitating Maintenance Decisions on the Dutch Railways Using Big Data: The ABA Case Study*, 2014 IEEE International Conference on Big Data (Big Data), IEEE.
- [23] Parida A., (2007), *Role of Condition Monitoring and Performance Measurements in Asset Productivity Enhancement*, 20th International Conference on Condition Monitoring and Diagnostic Engineering Management.
- [24] Jagadish H.V. et al., *Big Data And Its Technical Challenges*, "Communications of The ACM" 2014, t. 57, pp. 86-94.
- [25] Kumar U. et al., *Maintenance performance metrics: a state-of-the-art review*, "Journal of Quality in Maintenance Engineering" 2013, t. 19, pp. 233-277.
- [26] Orlowski W.J., Barley S.R., *Technology and institutions: what can research on information technology and research on organizations learn from each other?*, "MIS Quarterly" 2001, t. 25, pp. 145-165.
- [27] Simon P., (2013), *Too Big to Ignore: The Business Case for Big Data*, Wiley.
- [28] Chen G., Wua S., Wang Y., *The Evolvement of Big Data Systems: From the Perspective of an Information Security Application*, "Big Data Research" 2015, t. 2, pp. 65-73.
- [29] Fan J., Han F., Liu H., *Challenges of Big Data analysis*, "National Science Review" 2014, t. 1, pp. 293-314.
- [30] Hsieh J.C., Li A.H., Yang C.C., *Mobile, Cloud, and Big Data Computing: Contributions, Challenges, and New Directions in Telecardiology*, "International Journal of Environmental Research and Public Health" 2013, t. 10, pp. 6131-6153.
- [31] Bollen J., Mao H., Zeng X., *Twitter mood predicts the stock market*, "Journal of Computational Science" 2011, t. 2, pp. 1-8.

## THE ROLE AND POTENTIAL OF BIG DATA IN MAINTENANCE

**Abstract:** *Data collection applications are becoming increasingly important for production and maintenance processes. During daily operations, large amounts of files are generated and their abundance often paradoxically hinders the ability to extract useful knowledge. Moreover, the large amount of data stored in often unconnected databases makes it difficult to manually analyze valuable information for decision making. Nowadays, new tools and techniques are needed that can intelligently analyze collected information and provide useful knowledge for maintenance. This document explores the role of data for maintenance functions and situations where mismanaged data may cause problems instead of providing a solution.*

**Keywords:** maintenance, CMMS, resource management.