BUSINESS ANALYTICS SKILLS FOR THE FUTURE-PROOF CHAINS -

BUSINESS INTELLIGENCE

Autorzy:
Dario Šebalj
Dejan Mirčetić
Michał Adamczak

Dario Šebalj, Dejan Mirčetić, Michał Adamczak

BUSINESS INTELLIGENCE

Poznań 2025

Wydawca:
Wyższa Szkoła Logistyki
Estkowskiego 6
61-755 Poznań, Poland
www.wsl.com.pl

ISBN 978-83-62285-66-2 (online)

Recenenci:

- prof. Ljubica Milanović Glavan, University of Zagreb, Faculty of Economics and Business, Zagrzeb, Chorwacja
- prof. Igor Pihir, University of Zagreb, Faculty of Organization and Informatics, Varaždin, Chorwacja

Redakcja techniczna: Dario Šebalj, Josip Juraj Strossmayer University of Osijek, Faculty of Economics and Business in Osijek, Osijek, Chorwacja

Projekt okładki: Michał Adamczak, Wyższa Szkoła Logistyki, Poznań, Polska

Książka powstała w ramach projektu Business Analytics Skills for Future-proof Supply Chains (BAS4SC) [2022-1-PL01-KA220-HED-000088856] dofinansowanego ze środków programu ERASMUS+.

Przedmowa

W erze zdefiniowanej przez szybką transformację cyfrową i podejmowanie decyzji w oparciu o dane, umiejętność poruszania się, interpretowania i wykorzystywania ogromnych ilości danych nie jest już opcjonalna - jest niezbędna. Ta książka wyróżnia się jako aktualne i cenne źródło informacji dla studentów, nauczycieli i praktyków, którzy chcą rozwijać lub pogłębiać swoje kompetencje w zakresie analityki biznesowej, zwłaszcza w kontekście zarządzania łańcuchem dostaw i logistyką.

Opracowana w ramach współfinansowanego przez Erasmus+ projektu BAS4SC (Business Analytics Skills for Future-proof Supply Chains) książka odnosi się do zidentyfikowanej luki między programami nauczania, a wymaganiami przemysłu w całej Europie. Odzwierciedla ona rygorystyczny proces badawczy obejmujący analizę międzynarodowych programów studiów i bezpośrednie zaangażowanie ponad stu interesariuszy ze środowiska akademickiego i przemysłu. Rezultatem jest kompleksowy wybór najbardziej krytycznych umiejętności analitycznych potrzebnych przyszłym specjalistom ds. łańcucha dostaw.

Książka rozpoczyna się od podstaw zrozumienia i interpretacji danych, a następnie przechodzi przez podstawowe tematy, w tym analitykę danych biznesowych, eksplorację danych, uczenie maszynowe, zarządzanie procesami biznesowymi i systemy informacyjne. Szczególną uwagę poświęcono nowym obszarom, takim jak e-logistics, GIS i etyka danych - tematom, które w coraz większym stopniu kształtują konkurencyjne środowisko nowoczesnych łańcuchów dostaw.

Mamy nadzieję, że ta książka posłuży zarówno jako narzędzie do nauki, jak i profesjonalny przewodnik. Niezależnie od tego, czy jesteś studentem przygotowującym się do wejścia na rynek, czy praktykiem, który chce się dostosować i rozwijać w dynamicznym środowisku, przedstawione tu spostrzeżenia i umiejętności pozwolą ci podejmować mądrzejsze, szybsze i bardziej odpowiedzialne decyzje.

Dario Šebalj
Dejan Mirčetić
Michał Adamczak

SPIS TREŚCI

WPROWADZENIE

Niniejszy podręcznik, zatytułowany Business Intelligence, jest drugim z serii opracowanej w ramach projektu Business Analytics Skills for Future-proof Supply Chains (BAS4SC). W celu określenia treści tego podręcznika przeprowadzono kilka wstępnych działań badawczych. Po pierwsze, przeprowadzono kompleksowe badanie w celu przeanalizowania kursów analityki biznesowej w programach nauczania różnych programów studiów w Unii Europejskiej, Stanach Zjednoczonych i Wielkiej Brytanii, które dotyczą logistyki i zarządzania łańcuchem dostaw. Metodologia badania obejmuje przegląd publicznie dostępnych programów nauczania i opisów różnych programów studiów (studia licencjackie lub magisterskie). Badania koncentrowały się na programach studiów w dziedzinie ekonomii biznesu i nauk stosowanych. Analiza ta ujawniła lukę między wiedzą logistyczną i umiejętnościami analityki biznesowej wymaganymi w tej dziedzinie a tymi, które są obecnie oferowane studentom. Dzięki kompleksowym wywiadom i kwestionariuszom z kadrą dydaktyczną uniwersytetów, studentami i specjalistami z branży zidentyfikowano ponad 100 niezbędnych umiejętności analityki biznesowej. Stosując metodę klasyfikacji rankingowej ABC, 33 z tych umiejętności zostały wybrane do uwzględnienia w tej książce, z silnym naciskiem na obszary takie jak zarządzanie danymi, analiza procesów biznesowych, uczenie maszynowe, wizualizacja danych i systemy informacji biznesowej. Te wybrane umiejętności i zidentyfikowane potrzeby pozwoliły na stworzenie dziesięciu rozdziałów, z których każdy odnosi się do najważniejszych umiejętności wymaganych w tej dziedzinie.

Książka ta rozpoczyna się od położenia podwalin w Zrozumienie i interpretację danych, fundamentalnym rozdziale, który bada podstawowe koncepcje danych, takie jak typy danych, jakość i źródła. Wyjaśnia, w jaki sposób skuteczne podejmowanie decyzji jest zakorzenione w solidnym zrozumieniu danych, umożliwiając profesjonalistom wydobycie znaczących spostrzeżeń i uniknięcie uprzedzeń, które mogą zaciemnić osąd biznesowy. Tematy takie jak piramida danych-informacji-wiedzy-mądrości (DIWM), a także metody identyfikacji trendów i korelacji w danych biznesowych, stanowią solidną podstawę dla czytelników do poruszania się po złożonych krajobrazach danych.

Drugi rozdział, Business Data Analytics, kontekstualizuje duże zbiory danych w środowiskach biznesowych, wyjaśniając, w jaki sposób organizacje mogą wykorzystać analizę danych na dużą skalę w celu usprawnienia procesu decyzyjnego. Omawiając „pięć V”

big data - objętość, szybkość, różnorodność, prawdziwość i wartość - rozdział ten pomaga czytelnikom zrozumieć, w jaki sposób analityka big data wspiera cele operacyjne i strategiczne. Obejmuje on studia przypadków ilustrujące praktyczne zastosowania, pokazujące, w jaki sposób spostrzeżenia z dużych zbiorów danych mogą prowadzić do przewagi konkurencyjnej na dynamicznych rynkach.

W rozdziale poświęconym Eksploracji danych i odkrywaniu wiedzy czytelnicy zapoznają się z podstawowymi technikami wyodrębniania cennych wzorców z dużych zbiorów danych. Rozdział ten obejmuje koncepcje eksploracji danych, takie jak grupowanie, klasyfikacja, eksploracja reguł asocjacyjnych i wykrywanie anomalii, które pomagają firmom odkrywać przydatne informacje. Podkreślono kluczowe etapy procesu odkrywania wiedzy, od czyszczenia danych po interpretację wzorców, zapewniając, że spostrzeżenia uzyskane z eksploracji danych są zarówno istotne, jak i mają zastosowanie w podejmowaniu decyzji w świecie rzeczywistym.

Rozdział poświęcony Uczeniu Maszynowemu bada rolę uczenia maszynowego w analityce predykcyjnej i analityce biznesowej. Wyjaśnia podstawy uczenia maszynowego, w tym uczenia nadzorowanego i nienadzorowanego, oraz w jaki sposób techniki te pozwalają komputerom identyfikować wzorce, przewidywać i dostosowywać się w oparciu o nowe dane. Obejmując aplikacje uczenia maszynowego, takie jak systemy rekomendacji, prognozowanie popytu i wykrywanie anomalii, rozdział ten pokazuje, w jaki sposób algorytmy uczenia maszynowego umożliwiają organizacjom automatyzację podejmowania decyzji, poprawę obsługi klienta i optymalizację operacji.

Następnie, w rozdziale poświęconym Zarządzaniu procesami biznesowymi (BPM) i eksploracji procesów, przeanalizowano struktury i strategie stojące za wydajnymi procesami biznesowymi. W tej części opisano, w jaki sposób organizacje mogą analizować i udoskonalać swoje procesy, wykorzystując BPM i eksplorację procesów w celu dostosowania operacji do celów strategicznych. Rozdział zawiera narzędzia i ramy do oceny procesów biznesowych, wyjaśniając, w jaki sposób eksploracja procesów może odkryć wąskie gardła i nieefektywności, które utrudniają wydajność.

Przechodząc do systemów specjalistycznych, rozdział Systemy informatyczne w logistyce oferuje przegląd krytycznych systemów informatycznych, takich jak ERP (Planowanie Zasobów Przedsiębiorstwa), WMS (Systemy Zarządzania Magazynem) i TMS (Systemy Zarządzania Transportem). W rozdziale tym omówiono, w jaki sposób systemy te integrują różne funkcje logistyczne, usprawniają przepływ danych i poprawiają efektywność zarządzania zasobami w łańcuchach dostaw. Podkreślono znaczenie tych systemów

w dostarczaniu danych w czasie rzeczywistym i umożliwianiu elastycznego, opartego na danych zarządzania logistyką.

Rozdział poświęcony E-logistyce bada transformacyjny wpływ technologii cyfrowych na procesy logistyczne, podkreślając płynną integrację przepływów materialnych i cyfrowych. Rozpoczyna się od zdefiniowania e-logistyki i umiejscowienia jej w szerszej koncepcji e-biznesu, podkreślając, w jaki sposób procesy cyfrowe wspierają takie funkcje, jak pozyskiwanie materiałów, magazynowanie i transport, zwłaszcza w kontekście handlu elektronicznego. Rozdział przedstawia rozwój e-logistyki od wczesnych systemów planowania materiałowego do nowoczesnych rozwiązań ERP, WMS i TMS, które umożliwiają udostępnianie danych w czasie rzeczywistym w łańcuchach dostaw. Wraz z szybkim rozwojem gospodarki cyfrowej, e-logistyka stała się niezbędna dla organizacji dążących do poprawy pozycji konkurencyjnej i usprawnienia swoich operacji logistycznych.

Systemy informacji geograficznej (GIS) w logistyce zapewniają wgląd w analizę przestrzenną i jej rolę w optymalizacji sieci logistycznych. Rozdział ten opisuje, w jaki sposób aplikacje GIS umożliwiają firmom wizualizację, analizę i zarządzanie danymi geograficznymi, wspierając zadania takie jak optymalizacja tras, wybór lokalizacji i ocena ryzyka. Wykorzystując GIS, firmy mogą podejmować bardziej świadome decyzje, które uwzględniają czynniki geograficzne i środowiskowe, ostatecznie poprawiając wydajność i obsługę klienta.

Rozdział poświęcony metodom wizualizacji danych podkreśla znaczenie przekształcania złożonych danych w formaty wizualne, które ułatwiają zrozumienie i podejmowanie decyzji. Rozdział ten prowadzi czytelników przez skuteczne techniki wizualizacji, od prostych wykresów i grafów po zaawansowane interaktywne pulpity nawigacyjne. Omawia narzędzia i najlepsze praktyki, które sprawiają, że dane są dostępne dla interesariuszy, umożliwiając firmom jasne przekazywanie spostrzeżeń i prowadzenie świadomych działań.

Uwzględniając perspektywę etyczną, Rozdział zatytułowany Etyka danych i bezpieczeństwo informacji zajmuje się etycznymi rozważaniami na temat wykorzystania danych w kontekście biznesowym. W rozdziale tym omówiono przepisy dotyczące prywatności, praktyki ochrony danych oraz etyczne implikacje przetwarzania danych osobowych, w szczególności w analityce łańcucha dostaw. Omawiając tematy takie jak protokoły bezpieczeństwa danych i ramy etyczne, rozdział ten przygotowuje czytelników do przestrzegania standardów uczciwości i zaufania w środowiskach opartych na danych.

1. ZROZUMIENIE I INTERPRETACJA DANYCH

Autor: Dario Šebalj

W tym rozdziale wyjaśniono podstawowe koncepcje rozumienia i interpretacji danych. Dane stanowią podstawę skutecznego podejmowania decyzji i odgrywają kluczową rolę w napędzaniu sukcesu organizacji. Dzięki biegłości w analizie i interpretacji danych, osoby (analitycy danych, menedżerowie, przedsiębiorcy, entuzjaści danych itp.) mogą nabyć umiejętności niezbędnych do odnajdywania cennych spostrzeżeń z ogromnych ilości dostępnych informacji. Dane zapewniają analizę z faktycznym fundamentem. Umożliwiają organizacjom podejmowanie decyzji, które są bardziej obiektywne i oparte na faktach, umożliwiając im wyjście poza założenia i intuicję. Organizacje mogą wykorzystywać dane do znajdowania korelacji, trendów i wzorców, których w przeciwnym razie mogłyby nie zauważyć.

Znajdowanie nieefektywności organizacyjnych i obszarów do poprawy to kolejna korzyść analizy danych. Organizacje mogą znajdować wąskie gardła, usprawniać przepływy pracy i zwiększać ogólną wydajność procesów biznesowych poprzez ocenę danych operacyjnych. Ponadto analiza danych ułatwia ocenę sukcesu projektów lub strategii, które zostały wdrożone, co znacząco przyspiesza podejmowanie decyzji i usprawnia planowanie.

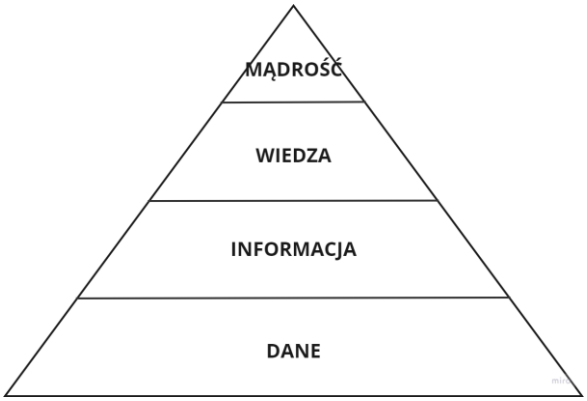
Procedura analizy danych musi być poprzedzona kluczowym opisaniem różnych typów i źródeł danych, omówieniem modelowania danych i podkreśleniem znaczenia jakości danych. Ten rozdział stworzy solidne podstawy Business Intelligence, zapewniając użytkownikom narzędzia do wykorzystania mocy płynącej z analizy danych i podejmowania świadomych decyzji.

1.1. Dane, informacje, wiedza, mądrość

Definiując termin dane, często jako punkt wyjścia stosuje się znaną piramidę danych-informacji-wiedzy-mądrości (DIWM) lub hierarchię DIWM. Rowley (2007) stwierdza, że hierarchia służy do identyfikowania i opisywania procesów zaangażowanych w transformację

jednostki na niższym poziomie hierarchii (takiego jak dane) w jednostkę na wyższym poziomie hierarchii (takiego jak informacje), a także kontekstualizowania danych, informacji, wiedzy

i okazjonalnie mądrości w odniesieniu do siebie nawzajem. Każdy poziom piramidy opiera się na poziomie poniżej, a aby podejmowanie decyzji oparte na danych było skuteczne, wymagane są wszystkie cztery poziomy (Cotton, 2023). Rysunek 1.1 przedstawia piramidę DIKW.



Rysunek 1.1 Piramida DIWM

Źródło: Rowley (2007)

Dane stanowią surowy materiał, który nie ma znaczenia. To treść, którą można bezpośrednio zaobserwować lub zweryfikować, niezorganizowany fakt, który jest wyrwany z kontekstu i trudny do zrozumienia (Brackett, 2015; Dalkir, 2023). Dane mogą mieć formę liczb, tekstu, obrazów itp. Bez interpretacji dane pozostaną bezsensowne. Przykład danych: zbiór danych zawierający odczyty temperatury zebrane ze stacji pogodowych.

Zbiór danych w kontekście, który jest istotny dla jednej lub więcej osób w danym momencie lub przez określony czas, nazywa się **informacją**. Są to przetworzone, zorganizowane, ustrukturyzowane i kontekstualizowane dane. Odpowiedzi na typowe pytania, takie jak „kto", „co", „gdzie" i „kiedy" można znaleźć w informacjach (Brackett, 2015; Cotton, 2023). Przykład informacji: Analizując dane dotyczące temperatury, można zauważyć, że średnia temperatura w poprzednim miesiącu jest wyższa niż w tym samym okresie w zeszłym roku.

Cotton (2023) twierdzi, że **wiedza** jest wynikiem analizy i interpretacji informacji, która ujawnia wzorce, trendy i powiązania. Oferuje wgląd w to, „jak" i „dlaczego" mają miejsce określone zdarzenia. Pociąga to za sobą głębsze zrozumienie podstawowych idei. Chaffey i Wood (2005, cytowani w Rowley, 2007) definiują wiedzę jako „połączenie danych i informacji, do których dodano opinię ekspertów, umiejętności i doświadczenie, co skutkuje cennym

zasobem, który może być wykorzystany do wspomagania podejmowania decyzji". Przykład wiedzy: Dzięki wiedzy uzyskanej z analizy historycznych danych dotyczących temperatury meteorolog może przewidzieć warunki pogodowe na nadchodzący tydzień.

Zdolność do zrozumienia podstawowych faktów i podejmowania świadomych decyzji oraz skutecznych działań jest znana jako **mądrość** (Cotton, 2023). Przykład mądrości: Korzystając z prognoz pogody i rozumiejąc lokalne warunki klimatyczne, rolnik może podjąć decyzję o zasianiu określonego rodzaju uprawy.

Zrozumienie wzajemnego oddziaływania danych, informacji, wiedzy i mądrości stanowi fundamentalną podstawę do wykorzystania potencjału Business Intelligence. W dalszej części tej książki zrozumienie to posłuży jako podstawa, w ramach której można wykorzystać moc danych, aby przekształcić je w praktyczne spostrzeżenia, kierując organizacje w stronę sukcesu w ciągle ewoluującym krajobrazie ery informacji.

1.2. Źródła danych i typy danych

Współcześnie dane są często określane jako „nowa ropa naftowa" – cenny zasób, który napędza innowacje i podejmowanie decyzji. W sercu każdego przedsięwzięcia opartego na danych leży spektrum typów danych, z których każde ma swoje unikalne cechy i znaczenie. Każdego dnia produkowana jest ogromna ilość danych. Obecne prognozy mówią, że na świecie jest 97 zettabajtów danych, a każdego dnia tworzonych jest ponad 2,5 kwintyliona bajtów danych. 90% danych na świecie wygenerowano w ciągu ostatnich dwóch lat (Marr, b.d.).

Według Kenetta & Shmueli (2016) „dane mogą pochodzić z różnych instrumentów gromadzenia: ankiet, testów laboratoryjnych, eksperymentów terenowych, eksperymentów komputerowych, symulacji, wyszukiwań w sieci, nagrań mobilnych itp. Dane mogą być pierwotne, zbierane na potrzeby badania, lub wtórne, zbierane z innego powodu. Dane mogą być jednowymiarowe lub wielowymiarowe, dyskretne, ciągłe lub mieszane. Dane mogą zawierać semantyczne niestrukturyzowane informacje w postaci tekstu, obrazów, dźwięku i wideo. Dane mogą mieć różne struktury, w tym dane przekrojowe, szeregi czasowe, dane panelowe, dane sieciowe, dane geograficzne itp. Dane mogą obejmować informacje z jednego źródła lub z wielu źródeł. Dane mogą mieć dowolny rozmiar i dowolny wymiar".

Dane są generowane szybciej i nieprzerwanie. Nowe dane są generowane i muszą być gdzieś przechowywane. Sieci społecznościowe, smartfony i technologia obrazowania wykorzystywana w diagnostyce medycznej to tylko kilka przykładów. Urządzenia i czujniki

automatycznie tworzą dane diagnostyczne, które muszą być analizowane i przechowywane natychmiast. Utrzymanie tej ogromnej ilości danych jest wystarczająco trudne, ale ich analiza w celu zidentyfikowania trendów i wydobycia cennych informacji jest o wiele trudniejsza, szczególnie gdy dane nie są zgodne z tradycyjnymi pojęciami struktury danych (EMC Education Services, 2015).

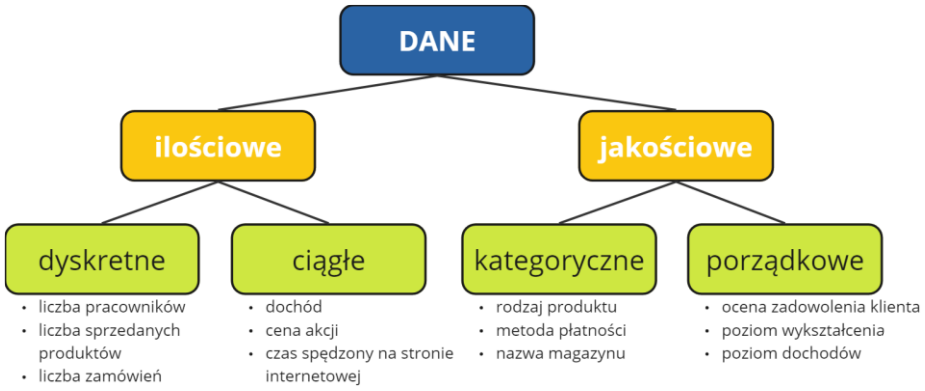
Według Blazqueza i Domenecha (2018) coraz częściej technologie związane z Internetem, smartfonami i inteligentnymi czujnikami są włączane do większości codziennych operacji biznesowych i osobistych. Na przykład wiele firm korzysta z mediów społecznościowych, aby promować swoje marki, oferować produkty online, używać smartfonów do śledzenia tras pokonywanych przez sprzedawców lub używać specjalistycznych czujników do rejestrowania pracy maszyn. Z drugiej strony ludzie używają komputerów, smartfonów i tabletów do przeglądania Internetu w poszukiwaniu produktów, komunikowania się ze znajomymi, dzielenia się opiniami i znajdowania drogi. Ponadto czujniki rozmieszczone w miastach, na autostradach i w miejscach publicznych, takich jak supermarkety, rejestrują codzienne ruchy i działania obywateli. Z tego powodu wszystkie te technologie produkują ogromną ilość nowych zdigitalizowanych i świeżych danych o działaniach ludzi i przedsiębiorstw. Po odpowiedniej analizie dane te mogą pomóc zidentyfikować trendy i śledzić skalę zachowań ekonomicznych, przemysłowych i społecznych. Dane te są ogromne i stale aktualizowane.

Sherman (2015) podkreśla, że problem może stanowić sytuacja, w której przedsiębiorstwo generuje więcej danych, niż jest w stanie zarządzać. Poprzez codzienne interakcje z klientami, partnerami i dostawcami firmy gromadzą ogromne ilości danych zarówno wewnętrznie, jak i zewnętrznie. Prowadzą badania rynku i śledzą informacje o swoich konkurentach. Witryny z kodami śledzącymi umożliwiają im śledzenie dokładnej liczby odwiedzających i ich pochodzenia. Przechowują i zarządzają danymi potrzebnymi do inicjatyw branżowych i wymogów rządowych. Obecnie istnieje Internet Rzeczy (IoT), który gromadzi dane z czujników wbudowanych w przedmioty ze świata rzeczywistego, takie jak obroże dla psów, rozruszniki serca i termostaty. To potop danych. Według EMC Education Services (2015) wśród źródeł Big Data o najszybszym tempie wzrostu znajdują się media społecznościowe i sekwencjonowanie genetyczne, które są przykładami wykorzystania niestandardowych źródeł danych do analizy.

Howson (2014) uważa, że sukces inicjatywy z zakresu Business Intelligence (BI) zależy od dostępności istotnych danych wysokiej jakości, obejmujących szeroki zakres źródeł danych niezbędnych do informowania procesów podejmowania decyzji.

Dane występują w różnych formach. Główne typy danych obejmują (Rysunek 1.2):

1. **Dane ilościowe (numeryczne)** – dane wyrażone w liczbach. Można je dalej podzielić na dane dyskretne i ciągłe. Dane dyskretne to dane, które mogą mieć tylko określoną wartość (np. liczba pracowników). Dane ciągłe mogą mieć nieskończoną liczbę możliwych wartości (np. cena produktu).
2. **Dane jakościowe** – ten typ danych można podzielić na dane kategoryczne i porządkowe. Dane kategoryczne reprezentują dane tekstowe, które można grupować w odrębne kategorie (np. miejsce urodzenia, region, kategoria produktu), podczas gdy dane porządkowe mogą być uszeregowane lub uporządkowane (np. satysfakcja klienta – bardzo niezadowolony, niezadowolony, neutralny, zadowolony, bardzo zadowolony; poziom wykształcenia – podstawowy, średni, licencjat, magister, doktorat).



Rysunek 1.2 Główne typy danych

Źródło: Opracowanie własne

Inną klasyfikacją danych stanowią dane ustrukturyzowane, półustrukturyzowane i nieustrukturyzowane.

Ustrukturyzowane dane mogą być przechowywane, przetwarzane i manipulowane w tradycyjnym systemie zarządzania relacyjną bazą danych. Dane te pochodzą z różnych źródeł, w tym strumieni kliknięć, formularzy internetowych, transakcji w punktach sprzedaży, czujników i maszyn. Mogą być wytwarzane przez ludzi lub maszyny. Mają one z góry określony format, rodzaj i organizację (EMC Education Services, 2015; Person & Porway, 2015).

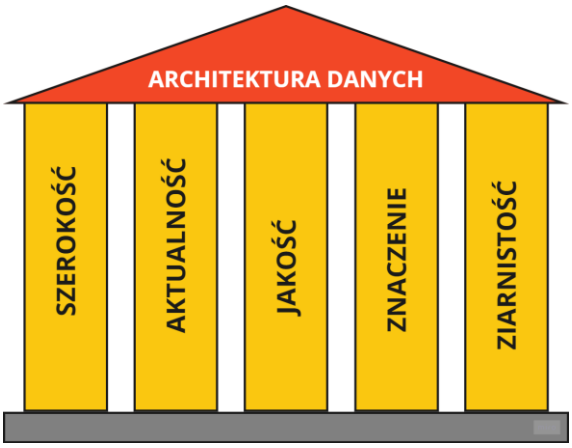
Dane półustrukturyzowane są organizowane za pomocą tagów, które pomagają nadać danym hierarchię i porządek, nawet jeśli nie pasują do ustrukturyzowanego systemu baz danych. Bazy danych i systemy plików często zawierają dane półustrukturyzowane. Pliki dziennika, tekst oznaczony tagami HTML, pliki XML i pliki danych JSON to niektóre z możliwych formatów przechowywania (McKinsey Global Institute, 2011; Person & Porway, 2015).

Ponieważ **niestrukturyzowane dane** są zazwyczaj wytwarzane przez działalność człowieka i nie mieszczą się w ustrukturyzowanym formacie bazy danych, są całkowicie niestrukturyzowane. Dokumenty tekstowe, pliki PDF, wpisy na blogu, wiadomości e-mail, zdjęcia i filmy są przykładami tego typu danych. (EMC Education Services, 2015; Person & Porway, 2015). Według Shermana (2015) niestrukturyzowane i półstrukturyzowane dane muszą być obsługiwane inaczej niż tradycyjne ustrukturyzowane dane.

Ta ogromna ilość ustrukturyzowanych i przede wszystkim nieustrukturyzowanych danych, które są generowane, gromadzone i przetwarzane z dużą prędkością i złożonością, nazywa się Big Data. McKinsey Global Institute (2011) definiuje Big Data jako „dane, których skala, dystrybucja, różnorodność i/lub terminowość wymagają użycia nowych architektur technicznych i analiz, aby umożliwić wgląd, który odblokowuje nowe źródła wartości biznesowej". Według Kitchena i McArdle'a (2016) w 2001 r. Doug Laney szczegółowo opisał, że Big Data charakteryzują trzy cechy (Trzy „V"):

- **objętość** (składająca się z ogromnych ilości danych),
- **prędkość** (tworzona w czasie rzeczywistym),
- **różnorodność** (ustrukturyzowana, półustrukturyzowana i nieustrukturyzowana).

Według Howsona (2014) podstawą udanej analizy biznesowej jest architektura danych (patrz Rysunek 1.3), która składa się z sześciu ważnych aspektów dotyczących danych: szerokości, terminowości, jakości, trafności i szczegółowości. Jakość danych jest filarem centralnym, ponieważ tak wiele wysiłku wkłada się w zapewnienie i poprawę jakości danych.



Rysunek 1.3 Architektura danych jako podstawa udanego BI

Źródło: opracowanie na podstawie Howson (2014)

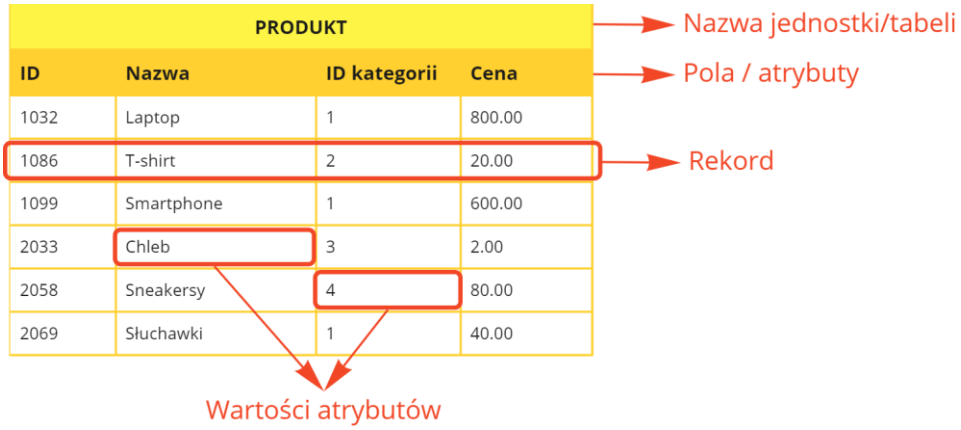
Jak już wspomniano, dane mogą być zbierane z różnych źródeł. Jest to związane z szerokością danych jako jednym z filarów architektury danych. Odnosi się to do możliwości uzyskania wielu źródeł danych, co jest obecnie, w erze Big Data, powszechnym sposobem zbierania danych. Z drugiej strony łączenie danych z wielu różnych systemów źródłowych również przyczynia się do problemów z jakością danych (Howson, 2014).

W następnym podrozdziale uwaga czytelnika zostanie przesunięta z pojęć związanych z rozumieniem danych na pojęcia związanie z wykorzystaniem ich potencjału. Zostaną zbadane relacyjne systemy zarządzania bazami danych (RDBMS) do przechowywania i pobierania danych, wraz z wizualną reprezentacją struktur baz danych za pomocą diagramów encji-związków (ER). Łącząc nasze zrozumienie źródeł danych i typów danych z modelowaniem i projektowaniem danych, zostanie wykonany znaczący krok w kierunku praktycznego zastosowania Business Intelligence.

1.3. Modelowanie i projektowanie danych

Model danych to formalna reprezentacja danych, z których system biznesowy korzysta i które generuje. (Dennis i in., 2018). Jak już wspomniano, ustrukturyzowane dane są zwykle przechowywane w relacyjnym systemie zarządzania bazą danych (RDBMS). Według Tilleya (2020) „system zarządzania bazą danych to zbiór narzędzi, funkcji i interfejsów, który umożliwia użytkownikom dodawanie, aktualizowanie, zarządzanie, uzyskiwanie dostępu i analizowanie danych". Niektóre popularne RDBMS to Oracle (Oracle), DB2 (IBM) i SQL Server (Microsoft).

W RDBMS dane są organizowane w tabelach, które zawierają zbiór rekordów przechowujących informacje o konkretnej jednostce. Tabele są reprezentowane jako dwuwymiarowe struktury z pionowymi kolumnami i poziomymi wierszami. Każda kolumna reprezentuje pole lub atrybut jednostki, podczas gdy każdy wiersz reprezentuje rekord, który jest wystąpieniem jednostki (Tilley, 2020). Rysunek 1.4 pokazuje przykład tabeli Product.



Rysunek 1.4 Przykład tabeli w RDBMS

Źródło: Opracowanie własne

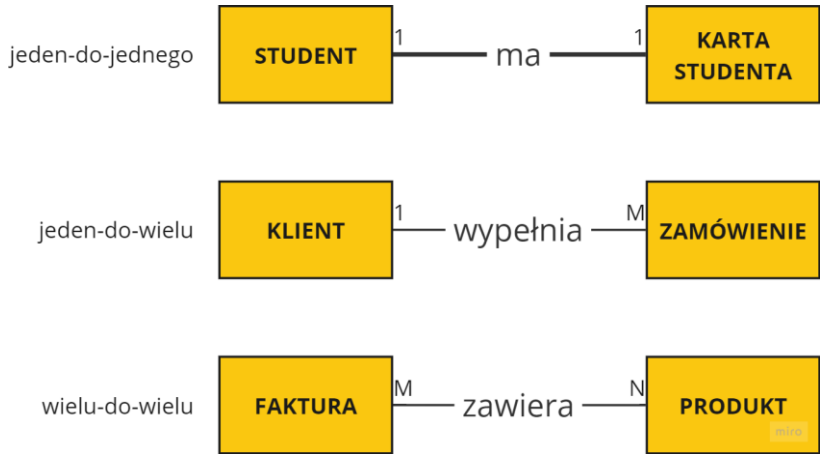
Tabela ta przedstawia prosty katalog produktowy zawierający 6 produktów, z których każdy ma unikalny identyfikator produktu, nazwę, kategorię i cenę.

Atrybut to szczególny rodzaj danych o jednostce. Na przykład identyfikator klienta, imię, nazwisko, adres, kod pocztowy, miasto, kraj, adres e-mail to atrybuty jednostki klienta. Wiersz w tabeli lub zbiór powiązanych pól opisujących pojedyncze wystąpienie jednostki (takiej jak klient) nazywa się **rekordem**.

Każda tabela w bazie danych musi mieć atrybut, który służy jako **klucz podstawowy**. Jest to pole (lub zestaw pól), które nadaje każdemu produktowi w tabeli odrębną tożsamość. Oznacza to, że w tabeli nie może być dwóch produktów o tym samym ID.

Tabele w bazie danych są często połączone z innymi tabelami w bazie danych, tzn. istnieje między nimi relacja. Relacje określają, w jaki sposób dane w jednej tabeli są powiązane z danymi w innej tabeli.

Według Tilleya (2020) między jednostkami mogą istnieć trzy typy relacji: jeden do jednego, jeden do wielu i wiele do wielu. Przykłady tych relacji pokazano na Rysunku 1.5.



Rysunek 1.5 Przykłady relacji pomiędzy jednostkami

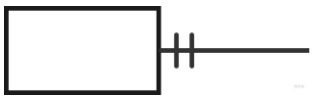
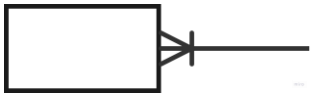
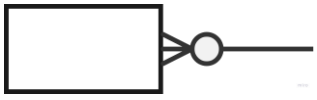
Źródło: Opracowanie własne

Logiczną strukturę bazy danych i relacje między tabelami można przedstawić wizualnie za pomocą **Diagramu Relacji Jednostki (ERD – ang. Entity Relationship Diagram)**.

Istnieją różne notacje do tworzenia ERD, z których najpowszechniejszymi są notacja Chen i notacja Martina („kruczej stopki"). W tej książce zostanie przedstawiona notacja Martina.

Zgodnie z notacją Martina („kruczej stopki"), jednostka (lub inaczej encja) jest reprezentowana przez prostokąt. Może to być osoba, miejsce, wydarzenie lub rzecz, na temat której zbierane są dane. Atrybuty są wymienione jako rzeczowniki w obrębie encji. Relacje między encjami są reprezentowane za pomocą linii, które łączą ze sobą encje. Relacje cechują się kardynalnością, która pokazuje, ile wystąpień jednej encji jest powiązanych z wystąpieniem innej (Dennis i in., 2018). W notacji „kruczej stopki" kardynalności są demonstrowane za pomocą różnych symboli. Na przykład pojedyncza kreska oznacza jeden, podwójna kreska oznacza jeden i tylko jeden, kółko oznacza zero, a „krucza stopka" oznacza wiele. Tabela 1.1. pokazuje różne symbole kardynalności i ich znaczenie.

Tabela 1.1 Przykłady kardynalności

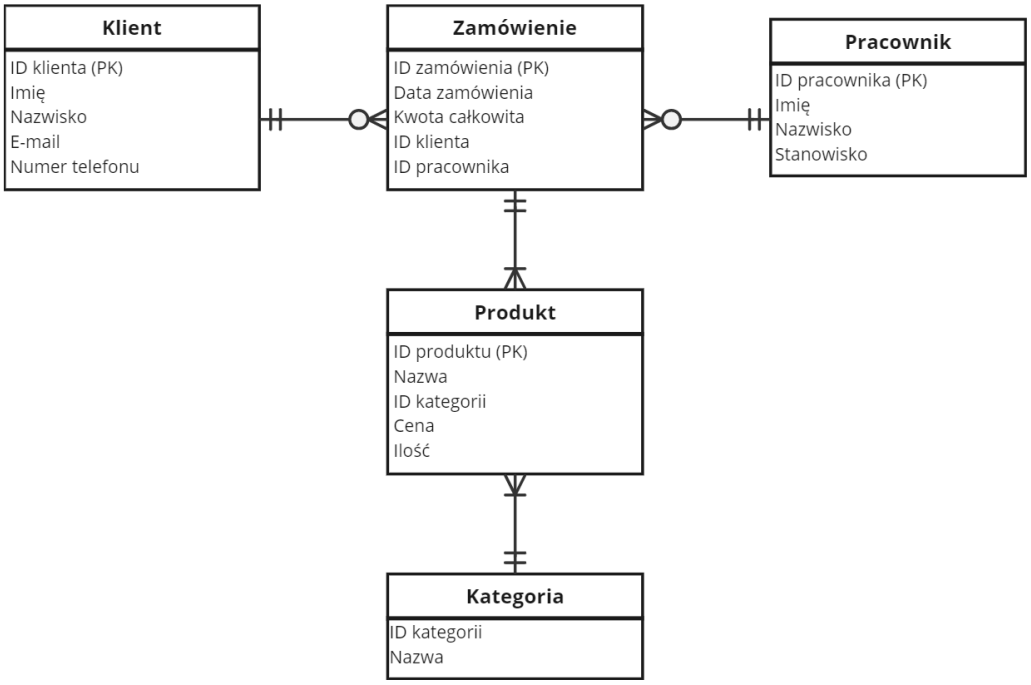
Symbol	Znaczenie
	Jeden i tylko jeden
	Jeden lub wiele
	Zero lub wiele

Symbol	Znaczenie
	Zero lub jeden

Źródło: Tilley (2020)

Według Dennisa i in. (2018) tworzenie ERD składa się z 3 etapów:

1. Zidentyfikuj jednostki.
2. Dodaj atrybuty i przypisz klucze podstawowe.
3. Zidentyfikuj relacje.



Rysunek 1.6 Przykład systemu sprzedaży ERD

Źródło: Opracowanie własne

Rysunek 1.6 przedstawia część systemu sprzedaży ERD. Każda jednostka w tym ERD jest przedstawiona jako prostokąt z listą jej atrybutów znajdującą się wewnątrz. Linie łączące jednostki służą do ukazania relacji jakie zachodzą pomiędzy nimi:

- Klient może złożyć wiele Zamówień. Każde zamówienie ma unikalnego klienta.
- Pojedyncze zamówienie może obejmować więcej niż jeden produkt. Każdy produkt jest powiązany z konkretnym zakupem.
- Zamówienie jest powiązane z Klientem i Pracownikiem, który je obsłużył.
- Produkt jest powiązany z Kategorią, ponieważ każdy produkt należy do pojedynczej kategorii.

Podstawowe zrozumienie modelowania i projektowania danych, jak pokazują ERD i RDBMS, jest niezbędne do efektywnego wykorzystania danych. To zrozumienie stanowi podstawę przejścia do pragmatycznego wdrażania danych poprzez podejmowanie decyzji opartych na danych, w którym systematycznie ustrukturyzowane modele danych dostarczają cennych spostrzeżeń motywujących świadome strategie biznesowe i procesy podejmowania decyzji.

1.4. Podejmowanie decyzji w oparciu o dane

Podejmowanie decyzji w oparciu o dane (DDDM – ang. Data-driven decision-making) to strategiczne podejście, które polega na analizowaniu i interpretowaniu danych w celu kierowania wyborami i działaniami. Wykorzystując spostrzeżenia pochodzące z ogromnych zestawów danych, firmy są w stanie precyzyjnie poruszać się w niepewnych obszarach, minimalizując w ten sposób ryzyko i maksymalizując możliwości. Nelson (2022) definiuje podejmowanie decyzji w oparciu o dane jako proces podejmowania strategicznych decyzji biznesowych, które współgrają z celami, zadaniami i inicjatywami organizacji, wykorzystując fakty, metryki i dane. Podejmowanie decyzji w oparciu o dane, według Provost & Fawcett (2013), to proces dokonywania wyborów, które opierają się bardziej na analizie danych niż na intuicji. Na przykład, specjalista ds. marketingu może wybierać reklamy, wykorzystując wyłącznie swoją rozległą wiedzę branżową i wyczucie tego, co spodoba się konsumentom. Alternatywnie, może oprzeć swoją decyzję na analizie danych pokazującej, jak klienci odbierają różne reklamy.

W tym podejściu decyzje nie są podejmowane na podstawie intuicji, ale twardych faktów. Proces decyzyjny polega na gromadzeniu, analizowaniu i interpretowaniu danych w celu identyfikacji wzorców, trendów i korelacji. Niezależnie od tego, czy chodzi o optymalizację wydajności operacyjnej, poprawę doświadczeń klientów czy udoskonalenie strategii produktów, decyzje oparte na danych umożliwiają firmom dostosowanie się do dynamicznie rozwijających się rynków.

Dzięki integracji danych z procesem podejmowania decyzji organizacje stają się bardziej elastyczne, responsywne i odporne na zmiany, wspierając tym samym innowacyjność i zrównoważony wzrost. Duża liczba prac badawczych wykazała, że podejmowanie decyzji w oparciu o dane wiąże się ze zwiększoną produktywnością (np. Brynjolfsson i McElheran, 2019; Sala i in., 2022; Colombari i in., 2023). Jest to powód, dla którego większość organizacji, zwłaszcza dużych, inwestuje w gromadzenie i analizowanie swoich danych. Ponad dwie trzecie

z ponad 300 dyrektorów objętych badaniem Bain & Company (2017) twierdzi, że ich firma inwestuje duże środki w analizę danych, podczas gdy ponad połowa przewiduje transformacyjne zwroty z inwestycji.

Istnieje pięć kroków podejmowania decyzji opartych na danych (Asana, 2022):

1. Zrozumienie wizji firmy.
2. Znajdowanie źródeł danych.
3. Czyszczenie i organizowanie danych.
4. Wykonywanie analizy danych.
5. Wyciąganie wniosków.

McKinsey Global Institute (2014) podaje, że organizacje oparte na danych mają 23 razy większą szansę na pozyskanie klientów, 6 razy większą szansę na utrzymanie klientów i 19 razy większą szansę na osiągnięcie rentowności.

Największe przedsiębiorstwa na świecie, które podejmują decyzje w oparciu o dane to Google, Amazon i Netflix.

Aby organizacja mogła w pełni wykorzystać potencjał Business Intelligence, kluczowe jest uwzględnienie jakości danych, która jest szczegółowo omówiona w poniższej sekcji. Jakość ta zapewnia precyzję i niezawodność wniosków oraz spostrzeżeń uzyskanych z danych.

1.5. Jakość danych

Stopień dokładności, spójności, niezawodności i przydatności do danego celu jest określany jako jakość danych. Jakość danych jest ważna w kontekście business intelligence i analizy danych, ponieważ wnioski i osądy wyciągnięte z danych zależą głównie od ich dokładności i niezawodności. Niska jakość danych może prowadzić do błędnych wniosków, wadliwych strategii i ostatecznie do szkodliwych wyników biznesowych. Według Gartnera (2021) każdego roku słaba jakość danych kosztuje organizacje średnio 12,9 miliona dolarów.

Jakość danych można scharakteryzować za pomocą sześciu najczęściej stosowanych wymiarów (Foote, 2022):

- **Dokładność:** jak dokładne są wartości atrybutów w danych?
- **Kompletność:** czy dane są kompletne, bez brakujących informacji?
- **Spójność:** jak spójne są wartości w bazach danych i między nimi?
- **Aktualność:** jak aktualne są dane?

- **Ważność:** w jaki sposób dane są zgodne z wstępnie zdefiniowanymi regułami biznesowymi?
- **Unikalność:** czy każdy rekord jest jednoznacznie zidentyfikowany, bez zbędnego przechowywania?

Według Shermana (2015) dane można uznać za wysokiej jakości, jeśli spełniają następujące kryteria (pięć „C" danych):

- **Czystość –** odnosi się do brakujących elementów, nieprawidłowych wpisów i innych podobnych problemów,
- **Spójność –** jednolitość i spójność danych w różnych źródłach i w samym zestawie danych,
- **Zgodność –** odnosi się do danych, które są zgodne z wstępnie zdefiniowanymi standardami i zasadami danych,
- **Aktualność –** niezbędne jest korzystanie z najnowszych danych do podejmowania decyzji i analiz,
- **Kompleksowość –** obejmuje wszystkie niezbędne elementy danych wymagane do zamierzonego procesu podejmowania decyzji, nie pomijając krytycznych informacji.

Według Kenetta i Shmueli (2016) niemal wszystkie dane muszą zostać oczyszczone, zanim będzie można ich użyć do dalszej analizy. Jednak cel determinuje stopień czystości i strategię czyszczenia danych. Wysokiej jakości informacje dla jednego celu i niskiej jakości informacje dla innego celu mogą znajdować się w tych samych danych.

Gartner (2023) niedawno zidentyfikował zestaw 12 działań mających na celu poprawę jakości danych. Działania te zostały sklasyfikowane w czterech odrębnych kategoriach, które należy wziąć pod uwagę przy ocenie integralności danych:

- Skup się na właściwych rzeczach, tworząc solidne fundamenty,
- Zastosuj odpowiedzialność za jakość danych,
- Ustal jakość danych „dostosowaną do celu",
- Zintegruj jakość danych z kulturą korporacyjną.

Według Howsona (2014) spójne, kompleksowe i dokładne dane są postrzegane jako mające wysoki stopień jakości. Trudno jest uzyskać dane wysokiej jakości, ponieważ problemy organizacyjne i własnościowe oddziałują na dane w bardzo dużym stopniu.

Umiejętność rozumienia i analizowania danych jest kluczowa dla podejmowania świadomych decyzji w dziedzinie business intelligence. Zrozumienie przejścia od surowych danych

do praktycznych spostrzeżeń, obejmujących różne źródła danych, typy, modelowanie i projektowanie, jest integralną częścią procesu czerpania wartości z informacji.

BIBLIOGRAFIA

1. Asana (2022). Data-driven decision making: A step-by-step guide [dostępne: https://asana.com/resources/data-driven-decision-making, dostęp 5 listopada 2023]
2. Bain & Company (2017). Closing the Results Gap in Advanced Analytics: Lessons from the Front Lines [dostępne: https://www.bain.com/insights/closing-the-results-gap-in-advanced-analytics-lessons-from-the-front-lines/, dostęp 5 listopada 2023]
3. Blazquez, D. & Domenech, J. (2018). Big Data sources and methods for social and economic analyses. Technological Forecasting & Social Change, 130, pp. 99-113.
4. Brackett, M. (2015). The Data-Information-Knowledge Cycle. Dataversity [dostępne: https://www.dataversity.net/the-data-information-knowledge-cycle/, dostęp 5 listopada 2023]
5. Brynjolfsson, E. & McElheran, K. (2019). Data in Action: Data-Driven Decision Making and Predictive Analytics in U.S. Manufacturing. Rotman School of Management Working Paper No. 3422397 [dostępne: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3422397, dostęp 5 listopada 2023]
6. Chaffey, D. & Wood, S. (2005). Business Information Management: Improving Performance using Information Systems. FT Prentice Hall.
7. Colombari, R., Geuna, A., Helper, S., Martins, R., Paolucci, E., Ricci, R. & Seamans, R. (2023). International Journal of Production Economics, 255.
8. Cotton, R. (2023). The Data-Information-Knowledge-Wisdom Pyramid. Datacamp [dostępne: https://www.datacamp.com/cheat-sheet/the-data-information-knowledge-wisdom-pyramid, dostęp 5 listopada 2023]
9. Dalkir, K. (2023). Knowledge Management in Theory and Practice, 4th Edition. MIT Press.
10. Dennis, A., Wixom, B. H. & Roth, R. M. (2018). Systems Analysis and Design, 7th Edition. Wiley.
11. EMC Education Services (2015). Data Science and Big Data Analytics: Discovering, Analyzing, Visualizing and Presenting Data. Wiley.

12. Foote, K. D. (2022). Data quality dimensions. Dataversity [dostępne: https://www.dataversity.net/data-quality-dimensions/, dostęp 5 listopada 2023]
13. Gartner (2021). How to Improve Your Data Quality [dostępne: https://www.gartner.com/smarterwithgartner/how-to-improve-your-data-quality, dostęp 5 listopada 2023]
14. Gartner (2023). Gartner Identifies 12 Actions to Improve Data Quality [dostępne: https://www.gartner.com/en/newsroom/press-releases/2023-05-22-gartner-identifies-12-actions-to-improve-data-quality, dostęp 5 listopada 2023]
15. Howson, C. (2014). Successful Business Intelligence: Unlock the Value of BI & Big Data, 2nd Edition. McGraw-Hill Education.
16. Kenett, R. S. & Shmueli, G. (2016). Information Quality: The Potential of Data and Analytics to Generate Knowledge. Wiley.
17. Kitchen, R. & McArdle, G. (2016). What makes Big Data, Big Data? Exploring the ontological characteristics of 26 datasets. Big Data & Society, 3(1).
18. Marr, B. (n.d.). How Much Data Do We Create Every Day? The Mind-Blowing Stats Everyone Should Read. Bernard Marr & Co. [dostępne: https://bernardmarr.com/how-much-data-do-we-create-every-day-the-mind-blowing-stats-everyone-should-read/, dostęp 5 listopada 2023]
19. McKinsey Global Institute (2011). Big data: The next frontier for innovation, competition, and productivity. [dostępne: https://www.mckinsey.com/~/media/mckinsey/business%20functions/mckinsey%20digital/our%20insights/big%20data%20the%20next%20frontier%20for%20innovation/mgi_big_data_full_report.pdf, dostęp 5 listopada 2023]
20. McKinsey Global Institute (2014). Five facts: How customer analytics boosts corporate performance [dostępne: https://www.mckinsey.com/capabilities/growth-marketing-and-sales/our-insights/five-facts-how-customer-analytics-boosts-corporate-performance, dostęp 5 listopada 2023]
21. Nelson, M. (2022). Beyond The Buzzword: What Does Data-Driven Decision-Making Really Mean?. Forbes [dostępne: https://www.forbes.com/sites/tableau/2022/09/23/beyond-the-buzzword-what-does-data-driven-decision-making-really-mean/?sh=2d35c4eb25d6, dostęp 5 listopada 2023]

22. Provost, F. & Fawcett, T. (2013). Data science and its relationship to big data and data-driven decision making. Big data, 1(1), pp. 51-59.
23. Rowley, J. (2007). The wisdom hierarchy: representations of the DIKW hierarchy. Journal of Information Science, 33(2), pp. 163–180.
24. Sala, R., Pirola, F., Pezzotta, G. & Cavalieri, S. (2022). Data-Driven Decision Making in Maintenance Service Delivery Process: A Case Study. Applied Sciences, 12(15).
25. Sherman, R. (2015). Business Intelligence Guidebook: From Data Integration to Analytics. Elsevier Inc.
26. Tilley, S. (2020). Systems Analysis and Design, 12th Edition. Cengage.

2. ANALITYKA DANYCH BIZNESOWYCH

Autor: Dejan Mirčetić

W erze digitalizacji, ze względu na ogromną ilość danych generowanych codziennie, nie ma możliwości wykorzystania tradycyjnej wiedzy i podejść do zarządzania procesami biznesowymi w różnych obszarach, a więc również do zarządzania logistyką i łańcuchami dostaw (Nikoličić i in. 2019). Web 2.0, wraz z Przemysłem 4.0, przetwarzaniem w chmurze, Internetem Rzeczy (ang. Internet of Things – loT), RFID (ang. Radio Frequency Identification) i innymi technologiami cyfrowymi doprowadziły do generowania, przechowywania i przesyłania dużych ilości danych. Wraz ze wzrostem objętości i złożoności danych rośnie również złożoność i czas potrzebny do analizy tych danych i wyciągania z nich wniosków.

Koncepcję Big Data po raz pierwszy wprowadzili Cox i Ellsworth w październiku 1997 r. w artykule biblioteki cyfrowej ACM (Tiwari i in. 2018). Badania nad pojęciem Big Data i jego konceptualizacja ewoluowały nieustannie. Początkowo Big Data charakteryzowała koncepcja 3V, która obejmowała **objętość, prędkość** i **różnorodność**, jak omówiono w poprzednim rozdziale. Następnie charakterystyka ta została rozszerzona o koncepcję 5V, obejmującą dwa dodatkowe atrybuty: **prawdziwość i wartość** (Nguyen i in. 2018; Tiwari i in. 2018). Objętość odnosi się do wielkości generowanych danych; objętość danych cyfrowych rośnie wykładniczo (Arunachalam i in. 2018). Różnorodność odnosi się do faktu, że dane mogą być generowane z heterogenicznych źródeł wewnętrznych i zewnętrznych, w formatach ustrukturyzowanych, półustrukturyzowanych i nieustrukturyzowanych (Nguyen i in. 2018). Prędkość odnosi się do szybkości generowania i dostarczania danych, które mogą być przetwarzane w partiach, w czasie rzeczywistym, prawie w czasie rzeczywistym lub w sposób uproszczony (Nguyen i in., 2018). Prawdziwość podkreśla znaczenie jakości danych, ponieważ wiele źródeł danych z natury zawiera pewien stopień niepewności i zawodności (Nguyen i in. 2018). Wartość odnosi się do znajdowania nowej wartości zawartej w danych, która może być wykorzystana do lepszego planowania biznesowego).

Analityka dużych zbiorów danych (ang. Big Data Analytics – BDA) obejmuje dwa wymiary: **Duże zbiory danych (ang. Big Data – BD)** opisane koncepcją 5V oraz **Analitykę biznesową (ang. Business Analytics – BA)**, która umożliwia uzyskanie wglądu w dane

poprzez zastosowanie statystyki, matematyki, ekonometrii, symulacji, optymalizacji lub innych technik wspomagających organizacje biznesowe w podejmowaniu lepszych decyzji (Wang i in. 2016). Big Data Analytics (BDA) obejmuje wykorzystanie zaawansowanych technik analitycznych służących wyodrębnieniu cennej wiedzy z ogromnych ilości danych o zmiennych typach celem wyciągnięcia wniosków poprzez odkrywanie ukrytych wzorców i korelacji, trendów i innych cennych informacji i wiedzy biznesowej. BDA pozwala na zwiększenie korzyści biznesowych, wydajności operacyjnej i eksplorację nowych rynków, a także stwarza możliwości rozwoju organizacji (Nguyen i in. 2018; Tiwari i in. 2018). BDA znajduje się w obszarze zainteresowania różnych środowisk zarówno akademickich, jak i biznesowych, szczególnie w logistyce i zarządzaniu łańcuchem dostaw

2.1. BDA w logistyce i zarządzaniu łańcuchem dostaw

Łańcuchy dostaw (ŁD; ang. Supply Chains – SC) stanowią sieć firm i zakładów zaangażowanych w przetwarzanie surowców w produkty końcowe, a także w dystrybucję produktów finalnych do klientów końcowych. W łańcuchach dostaw występują przepływy fizyczne, finansowe oraz informacyjne między różnymi firmami. Każdego dnia łańcuchy dostaw stają się coraz bardziej złożone, bardziej rozbudowane i bardziej globalne. Dlatego też, aby pomyślnie wdrożyć i zarządzać istniejącymi procesami w łańcuchach dostaw i ich ciągłym dostosowywaniem do warunków rynkowych, współczesne łańcuchy dostaw wymagają zatrudniania wysoko wykwalifikowanych ekspertów. Aby sprostać tym trudnym zadaniom, eksperci ŁD potrzebują formalnej edukacji, która zapewni im wiedzę i umiejętności z różnych dziedzin, przede wszystkim z logistyki, technologii informatycznych i ekonomii.

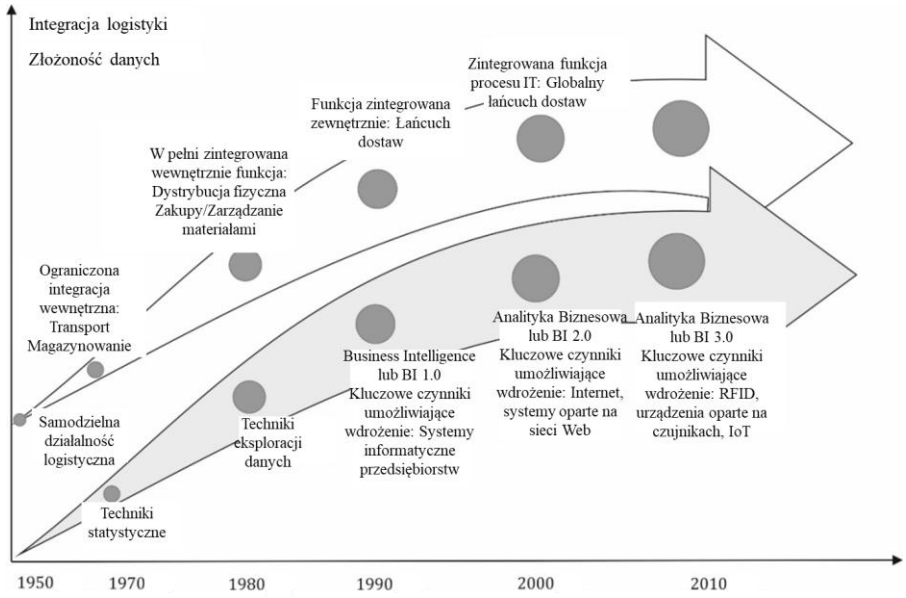
ŁD to zbiór elementów fizycznych, ich działań i procesów, poprzez które zachodzi ich interakcja. Elementy fizyczne, które tworzą strukturę łańcucha, stanowią stałą część ŁD. Decyzje dotyczące projektu struktury ŁD podejmowane są na poziomie strategicznym, a na poziomie taktycznym i operacyjnym podejmuje się decyzje dotyczące sposobów i zasad realizacji poszczególnych procesów logistycznych. Projektowanie stałego ŁD i zarządzanie poszczególnymi operacjami wspólnie stanowią zarządzanie łańcuchem dostaw, które definiuje wydajność łańcucha. Zgodnie z tym, ramy zarządzania ŁD składają się z trzech podstawowych elementów: (1) struktur ŁD; (2) procesów biznesowych; (3) i komponentów kontrolnych. Każdy z tych elementów jest bezpośrednio związany z celami ŁD, to znaczy ze stopniem spełnienia wymagań użytkowników końcowych, przy jednoczesnym poszanowaniu krytycznych wymiarów biznesu, które zależą od wydajności na rynku (kluczowe wskaźniki wydajności –

ang. Key Perfomance Inficators – KPI). We współczesnym świecie konkurencja nie odbywa się już między poszczególnymi organizacjami, ale między całymi łańcuchami dostawy. Efektywne zarządzanie ŁD stało się zatem potencjalnie cennym sposobem zapewniania przewagi konkurencyjnej i poprawy wydajności organizacji. Zarządzanie łańcuchem dostaw jest faktem biznesowym, a logistyka jest najpotężniejszym narzędziem umożliwiającym osiągnięcie ostatecznej przewagi strategicznej.

Firmy znajdują się pod silną presją poprawy planowania i wydajności ŁD z powodu takich czynników, jak rosnąca niepewność i konkurencja. Poprawa wydajności ŁD stała się ciągłym procesem, który wymaga analitycznego systemu pomiaru wydajności. Biorąc pod uwagę liczbę i różnorodność procesów logistycznych i procesów łańcucha dostaw zasoby wykorzystywane do ich realizacji, parametry, które je charakteryzują, jako podstawę do określania wydajności ŁD, wykorzystuje się dużą liczbę danych dotyczących aspektów: geograficznych, czasowych i ilościowych determinantów towarów, środków transportu, transportu – aktywów manipulacyjnych, pojemności magazynów, pracowników itp. Dane generowane poprzez wewnętrzne operacje, a także transakcje z dostawcami i klientami, mogą być wykorzystywane do odkrywania małych zmian, które mogą mieć duży wpływ na organizację pod względem wzrostu wydajności, a nawet oszczędności kosztów. Innymi słowy, ilość danych w każdym ŁD eksploduje z różnych źródeł danych, procesów biznesowych i systemów informatycznych. Wraz ze wzrostem ilości i złożoności danych rośnie również złożoność i czas potrzebny na analizę tych danych i uzyskanie z nich konkretnych informacji. Określanie, monitorowanie oraz poprawa logistyki i wydajności ŁD stają się bardziej złożone i obejmują wiele procesów, takich jak identyfikacja środków, definiowanie celów, planowanie, komunikacja, monitorowanie, raportowanie i sprzężenie zwrotne. W związku z tym konwencjonalne podejścia nie mogą być stosowane do podejmowania decyzji w ramach ŁD i zarządzania łańcuchem dostaw.

W zarządzaniu łańcuchem dostaw rośnie zainteresowanie analityką biznesową, zwaną również **Analityką Łańcucha Dostaw (ang. Supply Chain Analytics – SCA)**. W społecznościach biznesowych i akademickich SCA jest używane zamiennie z terminami takimi jak „analiza dużych zbiorów danych” i „analiza biznesowa” (Srinivasan i Swink, 2018). SCA odnosi się do wykorzystania danych oraz narzędzi i technik ilościowych w celu poprawy wydajności operacyjnej, często uwidocznionej przez takie wskaźniki, jak realizacja zamówień i elastyczność. Analityka w ŁD nie jest koniecznie nowym pomysłem, ponieważ różne techniki ilościowe i metody modelowania są od dawna stosowane w firmach produkcyjnych. Ostatni

wzrost zainteresowania analityką łańcucha dostaw wiąże się z nowymi wyzwaniami i możliwościami zarówno w środowiskach biznesowych, jak i informatycznych. Wyzwania te obejmują kwestie wynikające z zarządzania dużymi ilościami danych (np. dostępność i jakość danych) oraz radzenia sobie z niepewnością środowiskową. Prawidłowo stosowana SCA może mieć wpływ na kilka obszarów w ŁD i może generować znaczące korzyści w zakresie wydajności logistyki: usprawnione planowanie i harmonogramowanie; poprawiona responsywność; usprawnione planowanie popytu; optymalizacja zamówień; zoptymalizowane zarządzanie zapasami; poprawione planowanie uzupełniania zapasów. W ostatnich dekadach, na skutek rozwoju technologicznego, globalizacji i coraz bardziej wymagających klientów, zmieniły się również paradygmaty biznesowe. Na Rysunku 2.1 przedstawiono typowe okresy (z krótkim opisem) w ewolucji logistyki, SCM i BDA.



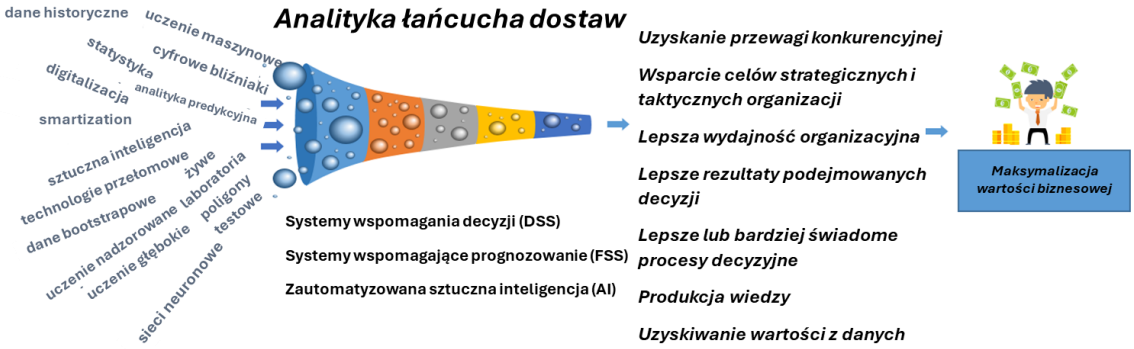
Rysunek 2.1 Ewolucja logistyki, SCM i BDA

Źródło: Arunachalam, Kumar & Kawalek (2018)

2.2. Narzędzia w Analityce Danych Biznesowych

Rysunek 2.2 przedstawia różne trendy, narzędzia i korzyści stosowane w BDA lub SCA. Wszystkie zaprezentowane techniki analityczne można podzielić na trzy typy: opisowe, predykcyjne i preskryptywne (ściśle powiązane z normami). **Analityka opisowa a**nalizuje dane i zdarzenia z przeszłości, aby określić jakie podejście zastosować w przyszłości. Poszukuje odpowiedzi co do przeszłych porażek i sukcesów. **Analityka predykcyjna** wykorzystuje dane historyczne, aby określić prawdopodobny przyszły wynik zdarzenia lub prawdopodobieństwo

wystąpienia sytuacji. Wykorzystuje wzorce znalezione w danych, aby zidentyfikować przyszłe ryzyka i możliwości. **Analityka preskryptywna** automatycznie syntetyzuje Big Data, reguły biznesowe i uczenie maszynowe, aby tworzyć przyszłe prognozy. Wykracza poza przewidywanie przyszłości, sugerując działania, które należy podjąć, aby osiągnąć pożądane cele. Jest ona również w stanie zademonstrować implikacje każdej możliwej decyzji i działać jako narzędzie wspomagające podejmowanie decyzji dla ekspertów ŁD. W poniższych podrozdziałach przedstawimy i opiszemy różne narzędzia analityczne stosowane w BDA w SCM. Ponadto skupimy się na strategiach zwiększania wiedzy w zakresie BDA dla ekspertów ŁD XXI wieku, poprzez kilka studiów przypadków.



Rysunek 2.2 Trendy, narzędzia i korzyści SCA

Źródło: Opracowanie własne

2.2.1. Analityka opisowa

Analityka opisowa zapewnia podsumowanie statystyk opisowych dla danej próbki danych, na przykład: średnia, moda, mediana, zakres, histogram i odchylenie standardowe. Analityka opisowa wskazuje, co wydarzyło się w przeszłości i wyprowadza informacje pochodzące ze znacznych ilości danych, aby odpowiedzieć na pytanie, co się dzieje w danej organizacji. Na podstawie informacji w czasie rzeczywistym o lokalizacjach i ilościach towarów w łańcuchu dostaw menedżerowie podejmują decyzje na poziomie operacyjnym (np. dostosowują harmonogram wysyłek, rozmieszczają pojazdy, składają zamówienia na uzupełnienie zapasów produktów itp.) (Souza, 2014). Analityka ta próbuje identyfikować szanse i problemy wykorzystując internetowy system przetwarzania analitycznego i narzędzia wizualizacji wspierane przez informacje dostarczane w czasie rzeczywistym, a także technologię raportowania (np. GPS, RFID, kod kreskowy transakcji). Typowymi przykładami analityki opisowej są raporty, które dostarczają historycznych spostrzeżeń dotyczących produkcji, finansów, operacji, sprzedaży, finansów, zapasów i klientów firmy (Tiwari i in. 2018).

2.2.2. Analityka predykcyjna

Analityka predykcyjna wykorzystuje dane historyczne do określenia prawdopodobnego przyszłego wyniku. Analityka predykcyjna w łańcuchach dostaw wyprowadza prognozy popytu z danych historycznych i odpowiada na pytania dotyczące tego, co się wydarzy lub co prawdopodobnie się wydarzy (Tiwari i in. 2018). Wykorzystuje sztuczną inteligencję, algorytmy optymalizacji i systemy eksperckie do przewidywania przyszłych zachowań na podstawie wzorców odkrytych w przeszłości, a także założenia, że zdarzenia historyczne się powtórzą. Wykorzystuje również wzorce zaszyte w danych do identyfikacji przyszłych ryzyk, możliwości optymalizacji oraz przewidywania przyszłości. Służy ona do uzupełniania brakujących informacji i eksplorowania wzorców danych za pomocą statystyki, symulacji i programowania.

2.2.3. Analityka preskryptywna

Analityka preskryptywna wyprowadza rekomendacje decyzyjne na podstawie modeli analityki opisowej i predykcyjnej, a także optymalizacji matematycznej, symulacji lub technik podejmowania decyzji wielokryterialnych. Wykracza poza przewidywanie przyszłych wyników, sugerując również działania, które przyniosą korzyści na podstawie przewidywań, pokazując decydentowi implikacje każdej opcji decyzyjnej. Analityka preskryptywna odpowiada na pytanie, co powinno się wydarzyć.

2.3. Otoczenie BDA

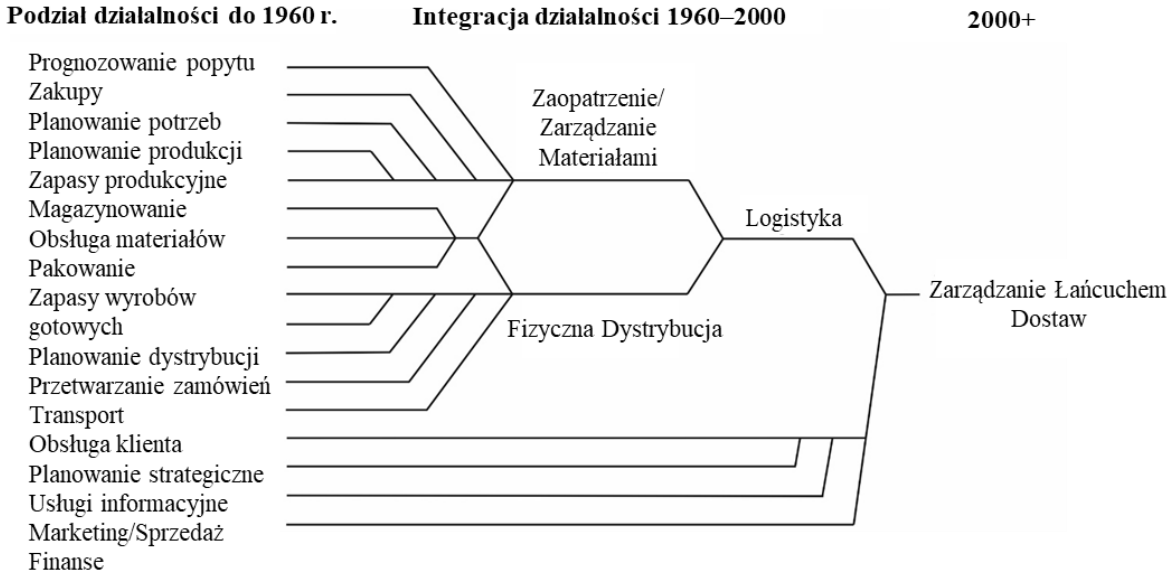
Głównym celem ekosystemu BDA jest dostarczanie wartości dla decydenta. W związku z tym podstawowym celem BDA jest udostępnienie wglądu w procesy biznesowe i dostarczenie podpowiedzi na temat tego, w jaki sposób obniżyć koszty i zwiększyć poziom obsługi dla klientów końcowych. Aby zrealizować swój cel, rozwiązania BDA są zwykle dostarczane w formie systemu wspomagania decyzji (ang. Decision Support System – DSS) lub systemu eksperckiego (ang. Exper System – ES). Dlatego w tym i kolejnych rozdziałach skupimy się na kluczowych filarach BDA: **danych biznesowych**, **eksploracji danych** i **odkrywaniu wiedzy** (analiza danych, DSS, platformy ES itp.).

Dane biznesowe

Koncepcja danych jest szczegółowo wyjaśniona w pierwszym rozdziale tej książki. Dane są kluczowym czynnikiem do przeprowadzania wszelkiego rodzaju analiz. **Dane biznesowe** są generowane w wyniku wykonywania procesów w danym środowisku biznesowym.

W przypadku ŁD istnieje wiele procesów i podprocesów zaangażowanych w dostarczanie usług klientowi ostatecznemu (Rysunek 2.3).

Rysunek 2.3 przedstawia strukturę ŁD, gdzie w przypadku SCA każdy z podanych procesów można obserwować jako generator danych biznesowych. Wygenerowane dane różnią się pod względem znaczenia i wpływu na ostateczne cele przedsiębiorstwa. W związku powyższym dane biznesowe można podzielić na dane **napędzane wewnętrznie** i **zewnętrznie**. Dane napędzane wewnętrznie to dane, które pojawiają się w wyniku struktury przedsiębiorstwa, hierarchii i sposobu, w jaki organizacja zdecydowała się działać (na przykład dane produkcyjne, dane dotyczące zasobów ludzkich, dane dotyczące dostaw, dane księgowe itp.). Dane te są różne dla każdego przedsiębiorstwa i służą do raportowania, analizowania i sprawozdawczości prawnej dla władz (dane księgowe). Interesującym jest również fakt, że firmy bezpośrednio kontrolują i wpływają na te dane, co czyni je istotnymi jedynie dla tej konkretnej organizacji.



Rysunek 2.3 Ewolucja logistyki i łańcuchów dostaw

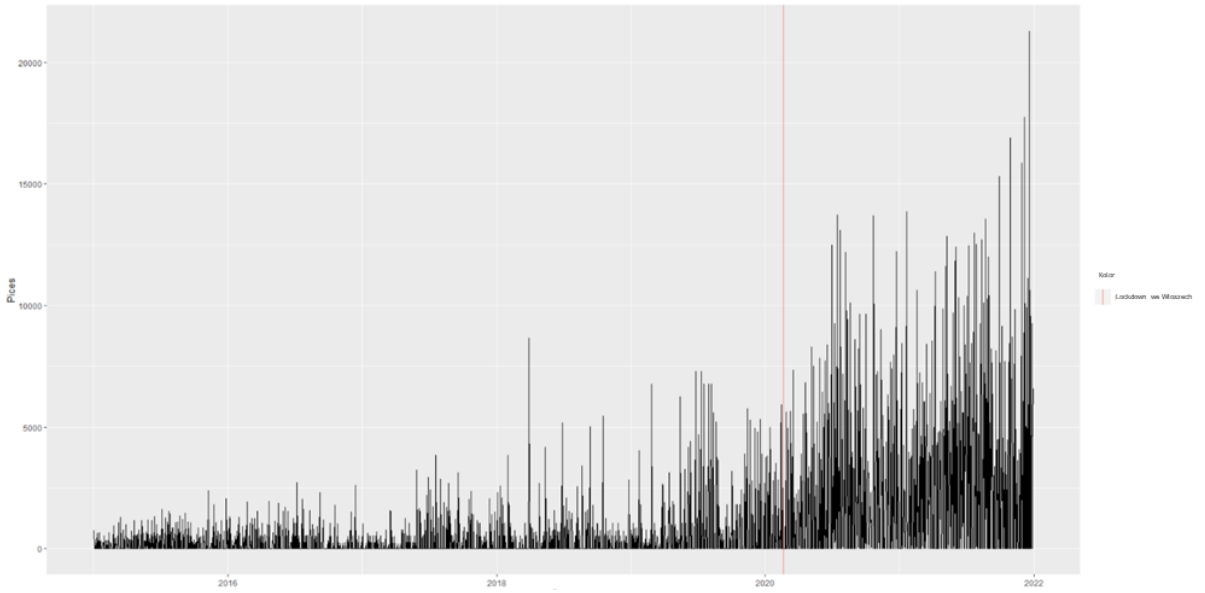
Źródło: Hesse & Rodrigue (2004)

Dane zewnętrzne odnoszą się do danych generowanych poza organizacją, które mogą być ogólnodostępne, niestrukturyzowane lub zbierane przez organizacje zewnętrzne (dane prywatne). Dla analityków ŁD szczególnie istotne są dane zewnętrzne udostępniane między ogniwami w łańcuchach dostaw, w tym dane o popycie rynkowym. Są one istotne, ponieważ stanowią wynik reakcji rynku na produkty i usługi firmy. Organizacja nie ma bezpośredniego wpływu na te dane, chociaż firmy próbują pośrednio poprawić reakcję rynkową klientów

wykorzystując procesy popytowe i ich podprocesy w postaci planowania popytu. Ponadto przedsiębiorstwa próbują modelować udział w rynku swoich produktów za pomocą działań planowania popytu, takich jak pakowanie, promowanie produktu, przeprowadzanie promocji sprzedaży, korzystanie z kilku kanałów dystrybucji itp.

Firmy inwestują dużo czasu, pieniędzy i wysiłku, aby lepiej zrozumieć i modelować swoje procesy zgodnie z danymi o popycie rynkowym. Z kilku powodów jest to bardzo trudne zadanie. Przede wszystkim firmy muszą określić infrastrukturę, procedury i umowy ze sprzedawcami detalicznymi, aby móc śledzić i rejestrować dane o popycie. Zazwyczaj firmy wykorzystują dane o sprzedaży gromadzone od partnerów na dole łańcucha (downstream) jako dane zastępcze dla danych o popycie. W rzeczywistości nie są to dane o popycie, a raczej dane o zakupach, które mogą znacznie zniekształcić informacje na temat popytu. Taka praktyka jest bardzo powszechna, ponieważ firmy niechętnie udostępniają swoje dane. Jednakże praktyka ta jest niewłaściwa, o czym duża część organizacji nie wie. Jedną z wad tego podejścia jest fakt, iż powoduje ono efekt byczego bicza (ang. bullwhip effect) wśród partnerów w ramach łańcucha dostaw. Innym podejściem do gromadzenia danych jest wykorzystanie danych z punktów sprzedaży (ang. point of sale – POS) sprzedawców detalicznych jako danych zastępczych dla danych popytowych (Syntetos i in., 2016). To podejście ma to również swoje zalety i wady, ponieważ nie bierze pod uwagę sytuacji braku towaru na półkach sklepowych. Takie działanie wymaga również stworzenia silnej infrastruktury informatycznej i umów ze sprzedawcami detalicznymi.

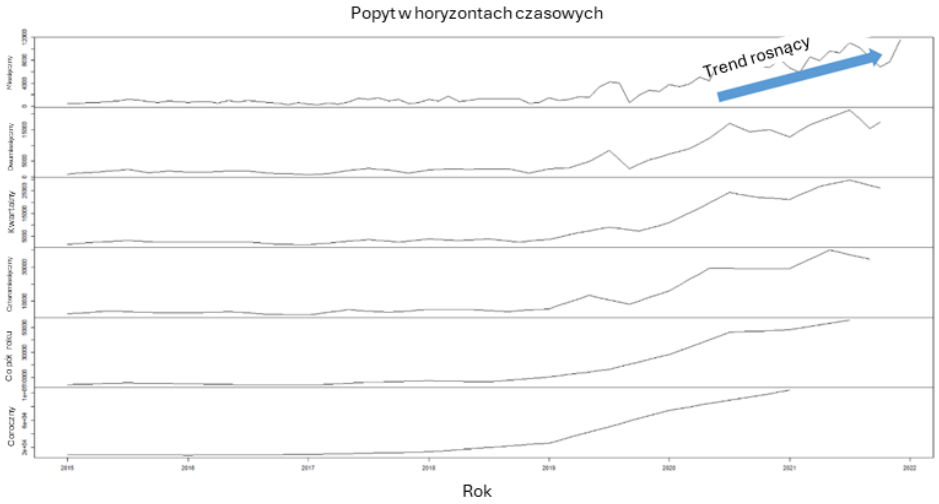
Drugim „problemem" wiążącym się z danymi generowanymi przez rynek jest fakt, że nie podążają one za zwykłymi procesami generowania statystyk. Stanowi to pewną trudność, ponieważ większość metod matematycznych i statystycznych zakłada, że dane podążają za pewnym procesem generowania statystyk. Jest to szczególnie zauważalne w ŁD, gdzie modele dotyczące zapasów zakładają, że popyt w okresie realizacji zamówienia podąża za rozkładem normalnym, a do obliczania zapasu bezpieczeństwa należy wykorzystać równania opierające się także o rozkład normlany. Według Mirčetića i in. (2017), Mirčetića i in. (2022) oraz Mirčetića i in. (2018), 90% danych z puli 97 serii w badaniu empirycznym przemysłu piwowarskiego nie podąża za rozkładem normalnym. Ponadto modele obliczania wielkości zapasów zakładają, że popyt jest deterministyczny i równomiernie rozłożony we wszystkich okresach, a w rzeczywistości tak nie jest. Na przykład na Rysunku 2.4 przedstawiono popyt na wstępnie pokrojone salami dla włoskiego producenta w okresie 2015-2022. Popyt wykazuje wyraźne zachowanie niedeterministyczne, tj. stochastyczne (z losowymi wahaniami i trendem).



Rysunek 2.4 Popyt na wstępnie pokrojone salami w plastrach w latach 2015-2022

Źródło: Opracowanie własne

Ponadto dzienne zapotrzebowanie wykazuje bardzo zmienne zachowanie, dlatego agregacja zapotrzebowania w różnych horyzontach czasowych (tygodniowym, miesięcznym itd.) charakteryzuje się wyraźnym trendem galopującego popytu (Rysunek 2.5).



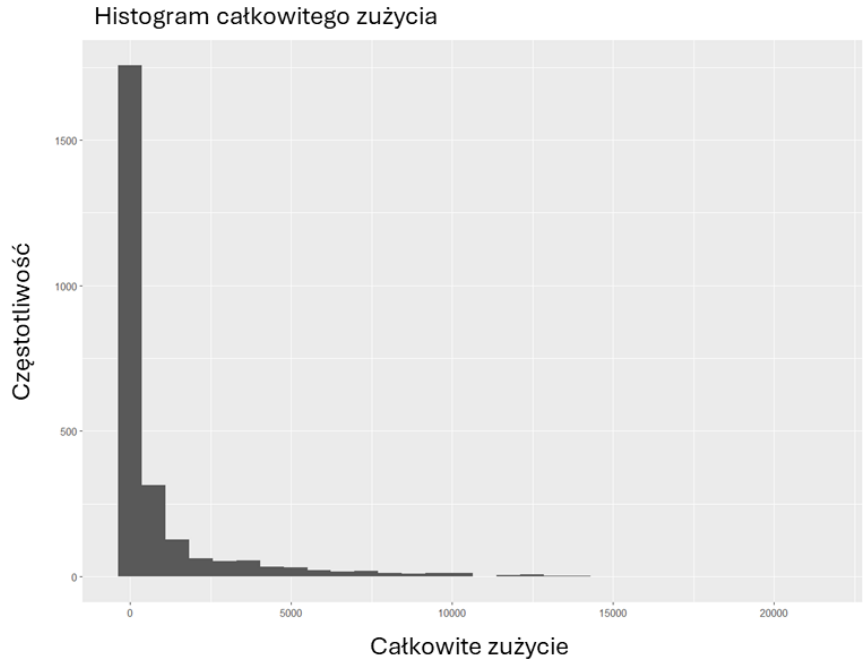
Rysunek 2.5 Agregacja popytu w różnych horyzontach czasowych

Źródło: Opracowanie własne

Agregacja popytu pokazuje wyraźną nową rzeczywistość od początku COVID-19 (wzrostowy trend w procesie popytu konsumpcyjnego)! To niezbity dowód na to, że popyt nie jest deterministyczny. Jeśli chodzi o założenie normalności, Rysunek 2.6 pokazuje znaczną

rozbieżność od rozkładu normalnego, mając na uwadze, że jest to rozkład skrajnie asymetryczny prawostronnie.

Rysunek 2.6 pokazuje, że popyt na salami uległ znacznej zmianie od początku pandemii COVID-19. Pierwszy lockdown we Włoszech miał miejsce 21.02.2020 (pionowa czerwona linia na Rysunku 2.4), po którym popyt na wstępnie pokrojone salami gwałtownie wzrósł i osiągnął historyczne maksimum. Popyt stale się zmieniał, osiągając najwyższy wzrost w dniu 20.12.2021, kiedy to sprzedano 21280 pokrojonych sztuk w ciągu jednego dnia.



Rysunek 2.6 Empiryczny rozkład popytu

Źródło: Opracowanie własne

Oprócz problemów z wyżej wymienionymi założeniami teoretycznymi dotyczącymi danych, obecna sytuacja w światowej gospodarce, a w konsekwencji w ramach łańcuchów dostaw, stawia przed analitykami biznesowymi kolejne pytania. Pytania te pojawiają się w wyniku pandemii, kryzysów wojennych, niedoborów zasobów, rosnącej inflacji, załamania światowych łańcuchów dostaw itp. Pytanie, które współcześni analitycy biznesowi muszą rozwiązać, zajmując się danymi wygenerowanymi po rozpoczęciu 2019 roku, brzmi: Jak długi okres i horyzont danych są teraz istotne dla obserwacji i modelowania? Widać to wyraźnie na Rysunku 2.4. Jeśli przyjrzymy się bliżej rysunkowi, zauważymy, że przed pandemią COVID-19 i lockdownem konsumenci nigdy nie spożywali wstępnie pokrojonego salami

w ilościach takich jak od pierwszego lockdownu. Można zauważyć, że spożycie danego produktu gwałtownie wzrosło. Teraz pojawia się kilka pytań:

- Czy to tylko moda na konsumpcję spowodowana specyficznymi warunkami podczas pandemii?
- Czy ten trend utrzyma się w przyszłości i firma będzie musiała zwiększyć produkcję?
- Czy dane sprzed COVID-19 (z lat 2015-2020) mają dziś jakąkolwiek wartość i czy należy je uwzględnić przy modelowaniu danych dotyczących popytu na wstępnie pokrojone salami?

Są to bardzo trudne pytania, na które nie da się odpowiedzieć bez odpowiedniej procedury analizy danych, która zostanie przedstawiona w kolejnych rozdziałach.

BIBLIOGRAFIA

1. Arunachalam, D., Kumar, N. & Kawalek, J. P. (2018). Understanding big data analytics capabilities in supply chain management: Unravelling the issues, challenges and implications for practice. Transportation Research Part E, 114, s. 416-436.
2. Hesse, M. & Rodrigue, J.-P. (2004). The transport geography of logistics and freight distribution. Journal of Transport Geography, 12(3), s. 171–184.
3. Mircetic, D., Rostami-Tabar, B., Nikolicic, S. & Maslaric, M. (2022). Forecasting hierarchical time series in supply chains: an empirical investigation. International Journal of Production Research, 60(8), pp. 2514-2533.
4. Mirčetić, D. (2018). Unapređenje top-down metodologije za hijerarhijsko prognoziranje logističkih zahteva u lancima snabdevanja (Doctoral dissertation), University of Novi Sad, Serbia.
5. Mirčetić, D., Nikoličić, S., Stojanović, Đ. & Maslarić, M. (2017). Modified top-down approach for hierarchical forecasting in a beverage supply chain. Transportation Research Procedia, 22, s. 193-202.
6. Mirčetić, D., Ralević, N., Nikoličić, S., Maslarić, M. & Stojanović, Đ. (2016). Expert system models for forecasting forklifts engagement in a warehouse loading operation: A case study. Promet-Traffic&Transportation, 28(4), s. 393-401.
7. Nikoličić, S., Maslarić, M., Mirčetić, D. & Artene, A. (2019). Towards more efficient logistic solutions: Supply chain analytics. In Proceedings of the 4th Logistics International Conference, Belgrade, Serbia (s. 23-25).
8. Souza, G. C. (2014). Supply chain analytics. Business Horizons, 57, s. 595-605.
9. Srinivasan, R. & Swink, M. (2018). An investigation of visibility and flexibility as complements to supply chain analytics: an organizational information processing theory perspective. Prod. Oper. Manag. 27(10), s. 1849–1867.
10. Syntetos, A. A., Babai, Z., Boylan, J. E., Kolassa, S. & Nikolopoulos, K. (2016). Supply chain forecasting: Theory, practice, their gap and the future. European Journal of Operational Research, 252(1), s. 1-26.
11. Tiwari, S., Wee, H. M. & Daryanto, Y. (2018). Big data analytics in supply chain management between 2010 and 2016: Insights to industries. Computers & Industrial Engineering, 115, s. 319-330.

12. Wang, G., Gunasekaran, A., Ngai, E. W. T. Papadopoulos Th. (2016). Big data analytics in logistics and supply chain management: Certain investigations for research and applications. Int. J. Production Economics, 176, s. 98-110.

13. Zhu, S., Song, J., Hazen, B. T., Lee, K. & Cegielski, C. (2018). How supply chain analytics enables operational supply chain transparency: An organizational information processing theory perspective. International Journal of Physical Distribution & Logistics Management, 48(1), s. 47-68.

3. EKSPLORACJA DANYCH (DATA MINING) I ODKRYWANIE WIEDZY (KNOWLEDGE DISCOVERY)

Autor: Dejan Mirčetić

Jak omówiono w Rozdziale 2, koncepcja ekstrakcji danych i generowania z nich wiedzy jest ściśle związana z **technikami Eksploracji Danych (ang. Data Mining techniques)**. We współczesnych przedsiębiorstwach techniki eksploracji danych wywierają duży wpływ na ogólną wydajność firm, ponieważ decyzje operacyjne, taktyczne i strategiczne są podejmowane w oparciu o informacje wejściowe uzyskane w procesie eksploracji danych. W Rozdziale 1 zauważono, że świat zawiera ponad 97 zettabajtów danych, a pojedyncze bazy danych często osiągają rozmiary terabajtów. W większości organizacji ilość danych podwaja się co dwa lata. Dane te są przechowywane na różnych platformach i w różnych formatach, zawierając dane ustrukturyzowane, nieustrukturyzowane i półustrukturyzowane (patrz Rozdział 1). Szacuje się, że do 90% danych biznesowych istnieje w formacie nieustrukturyzowanym. Ponadto znaczna część tych danych może zawierać błędy wynikające z niewłaściwego przechowywania lub formatowania czy też błędów ręcznych powstających podczas zbierania danych. W rezultacie nie wszystkie dane przechowywane przez organizacje cechują się wymaganą dokładnością lub wiarygodnością. Niemniej jednak ta ogromna ilość danych zawiera cenne informacje strategiczne dla firm. Jednak w obliczu tak dużej ilości złożonych typów danych pojawia się pytanie W jaki sposób można je skutecznie „wydobyć", aby uzyskać z nich istotne spostrzeżenia? Odpowiedź leży w eksploracji danych, która wspomaga zwiększanie przychodów i obniżanie kosztów poprzez szybkie i automatyczne wydobywanie przydatnej wiedzy i spostrzeżeń biznesowych z ogromnych zestawów danych.

Data Mining wyłoniło się z konieczności efektywnego wydobywania cennych informacji, wymagającego technik koncentrujących się na identyfikowaniu zrozumiałych wzorców, które można interpretować jako użyteczną lub interesującą wiedzę. Zatem **Data Mining jest iteracyjnym i interaktywnym procesem** mającym na celu odkrywanie ważnej, nowej, użytecznej i zrozumiałej wiedzy (wzorców, modeli, reguł itp.) w ogromnej bazie danych

(Behera i in., 2019). Głównym celem Data Mining jest **ujawnienie krytycznych spostrzeżeń**, które **wspierają podejmowanie decyzji** w organizacji biznesowej.

W kolejnych podrozdziałach zajmiemy się tym, w jaki sposób dane są „wydobywane" na potrzeby analityki biznesowej i Odkrywania Wiedzy (ang. Knowledge Discovery).

3.1. Czym jest Eksploracja Danych (Data Mining)?

Przed zdefiniowaniem Data Mining ważne jest, aby umieścić to pojęcie w kontekście terminów, z którymi jest powszechnie kojarzone i powiązane oraz z tymi, z którymi jest często błędnie utożsamiany. Osoby niebędące ekspertami często mylą terminy Data Mining i Big Data technology. Są to jednak dwa odrębne pojęcia. **Big Data opisuje** niezwykle duże i złożone **zestawy danych**, które wymagają specjalistycznych aplikacji i oprogramowania do przetwarzania. Z drugiej strony **Data Mining idzie o krok dalej**, obejmuje analizę tak ogromnych ilości danych w celu odkrycia ukrytych reguł i wzorców, które mogą nie być widoczne na pierwszy rzut oka.

Data Mining to szerokie pojęcie obejmujące różne techniki analityczne, w tym statystykę, sztuczną inteligencję i uczenie maszynowe. Metody te służą do przeszukiwania ogromnych ilości danych przechowywanych w bazach danych organizacji lub repozytoriach online. Głównym celem eksploracji danych jest odkrywanie wzorców w zestawie danych. **Analityka Biznesowa (ang. Business Analytics – BA)** odnosi się do kompleksowego procesu wykorzystywania umiejętności, technologii, ustalonych praktyk i algorytmów związanych z Data Mining. **Stąd Data Mining powszechnie służy jako zaplecze funkcji BA**, podczas gdy front-end funkcji BA składa się z metryk raportowania dla kadry kierowniczej i zestawionych informacji przedstawionych w formacie, który umożliwia menedżerom podejmowanie świadomych decyzji biznesowych. Podczas korzystania z Data Mining profesjonaliści BA działają jak „detektywi danych" (Lee, 2013), analizując dane celem lepszego opisania i zrozumienia obecnej i przeszłej sytuacji organizacji (analityka opisowa), przewidzenia przyszłych wyników (analityka predykcyjna) i podjęcia skutecznych działań (analityka preskryptywna).

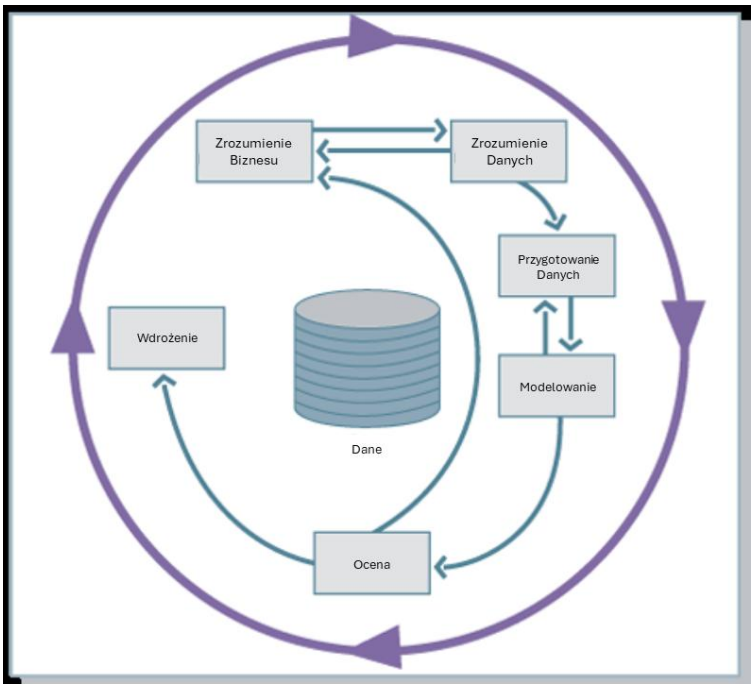
Data Mining jest podstawowym elementem procesu **Odkrywania Wiedzy w Bazach Danych (ang. Knowledge Discovery in Databases – KDD)**, ale stanowi tylko jeden krok w całym procesie. Aspekt Data Mining procesu KDD koncentruje się na wykorzystaniu algorytmów do ekstrakcji i identyfikacji wzorców z danych. W szerszym procesie KDD te

wydobyte wzorce są oceniane i potencjalnie interpretowane w celu ustalenia, które wzorce można uznać za nową „wiedzę" (Behera i in., 2019). Zdefiniowane w ten sposób, z wykorzystaniem Data Mining jako zapleczem i Knolwedge Discovery jako swego rodzaju front-endem BA, wzorce te reprezentują, wraz z danymi biznesowymi, kluczowe filary analityki biznesowej, jak opisano w Rozdziale 2.

Data Mining wykorzystuje różnorodne algorytmy do eksploracji ogromnych zbiorów danych, identyfikując wzorce, które mogą dostarczać cenne informacje biznesowe. Data Mining to narzędzie, a nie magiczne rozwiązanie. Nie obserwuje biernie bazy danych organizacji i nie ostrzega jej zarządców o interesujących wzorcach. Zrozumienie danego przedsiębiorstwa, jego danych i metod analitycznych pozostaje kluczowe. Data Mining pomaga analitykom biznesowym odkrywać wzorce i relacje w danych, ale nie określa wartości tych wzorców dla organizacji. Dlatego Data Mining nie zastępuje wykwalifikowanych analityków biznesowych, zamiast tego zapewnia im potężne nowe narzędzie do usprawnienia ich pracy.

Data Mining obejmuje proces obliczeniowy polegający na identyfikowaniu trendów, reguł, ukrytych wzorców i innych cennych informacji poprzez analizę dużych zestawów danych. Data Mining dostosowuje swoją technikę korzystając z wielu obszarów badawczych, w tym statystyki, uczenia maszynowego, systemów baz danych, wizualizacji, sieci neuronowych itp. Jest to proces wydobywania praktycznej wiedzy z różnych źródeł danych przechowywanych w zróżnicowanych formatach. Data Mining stał się coraz bardziej istotny w ostatnich latach ze względu na postęp obserwowany w technologiach przechowywania danych (ang. Big Data), sztucznej inteligencji (ang. Artificial Intelligence – AI) i automatyzacji procesów robotycznych (ang. Robotic Process Automation – RPA).

Proces Odkrywania Wiedzy w Bazach Danych (KDD) obejmuje wykorzystanie bazy danych, w tym dokonania niezbędnego wyboru, wstępnego przetwarzania, wstępnego próbkowania i transformacji, w celu zastosowania algorytmów Data Mining do identyfikacji wzorców (Behera i in., 2019). Obejmuje on również ocenę wyników Data Mining. Wspólnym standardem opisu kroków procesu Knowledge Discovery jest Międzybranżowy Standardowy Proces Eksploracji Danych (ang. Cross-Industry Standard Process for Data Mining – CRISP-DM), który pokazano na Rysunku 3.1.



Rysunek 3.1 Międzybranżowy Standardowy Proces Eksploracji Danych (ang. Cross-Industry Standard Process for Data Mining – CRISP-DM)

Źródło: Rahman i in. (2016)

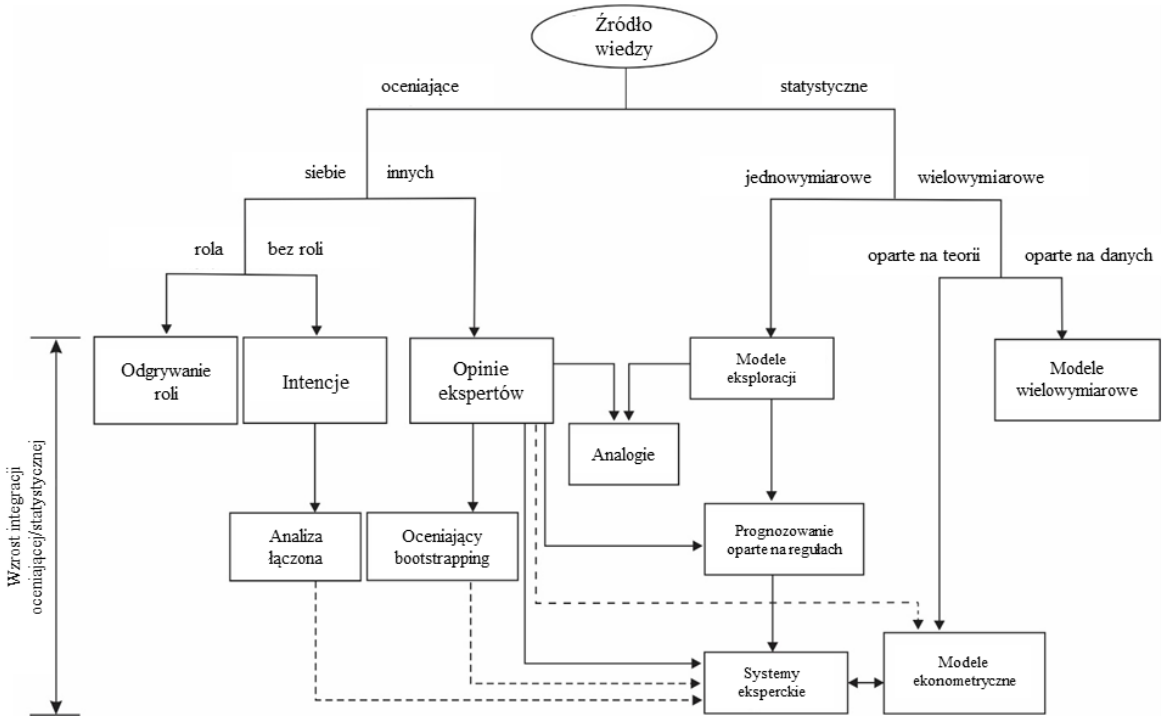
Na Rysunku 3.1 pierwszą fazę procesu stanowi zrozumienie biznesowe, obejmujące zrozumienie celów, które mają zostać osiągnięte, a także przeprowadzenie szczegółowego ustalania faktów na temat zasobów i założeń. Druga faza, zrozumienie danych, skupia się na badaniu różnych opisowych charakterystyk danych. Trzecia faza, przygotowanie danych, jest najbardziej wymagającą i czasochłonną częścią procesu KDD, mającą na celu wybranie odpowiednich danych i odpowiednie sformatowanie ich do analizy. Ta faza obejmuje działania takie jak wybór danych, filtrowanie, transformacja i integracja. Czwarta faza, modelowanie, obejmuje stosowanie metod analitycznych i wybór najbardziej odpowiednich algorytmów. Ta faza obejmuje również weryfikację jakości modelu poprzez testowanie i walidację krzyżową (sprawdzian krzyżowy). Piąta faza, ocena, obejmuje interpretację i ocenę odkrytej wiedzy (Rahman i in., 2016).

3.2. Odkrywanie Wiedzy w Logistyce i Zarządzaniu Łańcuchem Dostaw

W kontekście ŁD i logistyki Data Mining różni się od innych aplikacji ogólnego przeznaczenia. Powód jest związany z wspomnianą i omówioną wcześniej różnorodnością źródeł i struktury danych biznesowych, które stanowią duże wyzwanie dla inżynierów

projektujących odpowiednie rozwiązania modelowe. Proces generowania wiedzy z danych można podsumować na Rysunku 3.2.

Na podstawie Rysunku 3.2, odkrywanie i generowanie wiedzy z danych jest w dużym stopniu zależne od źródła wiedzy i można je podzielić na **oceniające** i **statystyczne**. Oba te kierunki mają swoje zalety, ale procedura ekstrakcji wiedzy ze źródła jest zasadniczo różna. Aby zastosować bardziej formalne podejście do eksploracji danych, tj. podejście statystyczne, kluczowym fundamentem jest istnienie danych ilościowych. W łańcuchach dostaw istnieje duża liczba sektorów, lokalizacji i transakcji, generujących przepływy danych, które można wykorzystać do eksploracji i ekstrakcji przydatnych informacji zwrotnych. Z drugiej strony, w ŁD i logistyce istnieje również wiele źródeł danych, które nie mają charakteru ilościowego, a zatem nie podlegają formalnym procedurom ilościowym. Zamiast tego dane te są tradycyjnie poddawane eksperckim panelom oceniającym (metoda delficka), a decyzje są podejmowane na podstawie doświadczenia, wiedzy i autorytetu ekspertów.



Rysunek 3.2 Zasady ekstrakcji wiedzy z danych

Źródło: Armstrong (Ed.) (2001)

W tej książce nacisk zostanie położony na **techniki ilościowe/matematyczne**, choć krótko omówimy również niektóre metodologie przechwytywania wiedzy eksperckiej w ustrukturyzowanych ramach, tj. omówimy System Eksperckie (ang. Expert Systemes – ES) i ich zastosowania oraz przykłady.

3.3. Podejście Delfickie do Oceniającego Tworzenia Wiedzy

Cenne spostrzeżenia doświadczonych profesjonalistów w dziedzinie łańcucha dostaw i logistyki często pozostają niezauważone. Te spostrzeżenia należy udostępniać mniej doświadczonym profesjonalistom w tej dziedzinie. Metoda delficka jest jednym ze sposobów pozyskiwania i rozpowszechniania wiedzy eksperckiej. Według (Steurer, 2011) metoda delficka, nazwana na cześć wyroczni delfickiej, która początkowo była wykorzystywana do konsultacji w różnych sprawach publicznych i osobistych w starożytnej Grecji, w latach 50. XX wieku przekształciła się w technikę, w której eksperci zastępowali wyrocznię, aby osiągnąć konsensus wśród grupy ekspertów w danej dziedzinie. „Project Delphi", finansowany przez Siły Powietrzne USA, był pierwszym projektem wykorzystującym tę metodę do prognozowania rozwoju technologicznego. Od tego czasu metoda delficka ewoluowała i udoskonalała się, znajdując zastosowania w różnych dyscyplinach naukowych. Metoda delficka została opracowana w celu osiągnięcia wiarygodnego konsensusu ekspertów, często służąc jako substytut dowodów empirycznych, gdy takich dowodów brakuje. Oznacza to, że technika delficka jest procesem iteracyjnym, w którym eksperci anonimowo wydają osądy na konkretny temat, mając na celu osiągnięcie konsensusu i sprzeciwu wraz z ich uzasadnieniami. Jest to wysoce ustrukturyzowany proces komunikacji grupowej, w którym eksperci oceniają niepewną i niekompletną wiedzę (Naisola-Ruiter, 2022). Według Paivarinty i in. (2011) metoda delficka, między innymi, jest szeroko wykorzystywana w badaniach nad systemami informacyjnymi. Jest stosowana do wybierania projektów w zakresie Systemów Informacyjnych (ang. Information Systems – IS), ustalania priorytetów ryzyka projektów rozwoju oprogramowania, definiowania wymagań projektu IS, określania kluczowych kwestii w zarządzaniu IS, tworzenia ram dla działań związanych z manipulacją wiedzą, zrozumienia ról i zakresu systemów zarządzania wiedzą w organizacjach oraz monitorowania badań nad systemami informacyjnymi w zakresie offshoringu (przeniesienia wybranych procesów działalności poza granicę kraju).

Metoda delficka stała się standardową praktyką w zakresie ilościowego określania wyników procesów pozyskiwania uwagi w grupie. Jest wykorzystywana w różnych dyscyplinach do prognozowania trendów, ustalania priorytetów obszarów badawczych, oceny potencjalnych skutków różnych wyborów politycznych, ustalania wskaźników wydajności i opracowywania wytycznych klinicznych. **Techniki delfickie są również stosowane w obszarze ŁD i logistyki**. Na przykład jest ona wysoce zalecana jako instrument do identyfikacji i oceny ryzyka dostaw, do różnego rodzaju oceny w procesach logistycznych, do określania najlepszych

praktyk logistycznych, do strategicznego podejmowania decyzji i opracowywania polityki przedsiębiorstwa, do mapowania przyszłych praktyk SCM i do prognozowania logistyki. Cztery kluczowe cechy lub podstawowe zasady metody delfickiej to:

- proces iteracyjny i wieloetapowy (oraz zbieranie danych);
- opinie uczestników (kontrolowane na pewnym poziomie) z możliwością poprawienia swoich odpowiedzi'
- statystyczne określenie odpowiedzi grupy;
- pewien stopień anonimowości.

Typowy proces delficki obejmuje przedstawienie serii pytań w wielu rundach. Paneliści, wybrani ze względu na swoją wiedzę i doświadczenie, odpowiadają anonimowo. Po każdej rundzie następuje informacja zwrotna na temat zagregowanych odpowiedzi, pozwalając uczestnikom dostrzec, jak kształtują się ich odpowiedzi w kontekście odpowiedzi całego panelu. Paneliści mogą następnie dostosować swoje odpowiedzi i podać uzasadnienie wszelkich zmian w kolejnych rundach. Ten iteracyjny proces trwa do momentu osiągnięcia konsensusu lub ukończenia ustalonej liczby rund.

Kroki przeprowadzania metody delfickiej

Metoda delficka to ustrukturyzowane podejście, które polega na zbieraniu opinii i spostrzeżeń ekspertów w celu osiągnięcia konsensusu w danym temacie. Proces ten zazwyczaj obejmuje cztery główne kroki (Rysunek 3.3).

Rysunek 3.3 Kroki przeprowadzania metody delfickiej

Źródło: Steurer (2011)

Krok 1 – Określenie celów: Pierwszy krok obejmuje określenie celów i zakresu badania delfickiego. Krok ten zawiera więc identyfikację konkretnych pytań lub tematów wymagających wkładu eksperta i nakreślenie kluczowych kwestii do omówienia. Ten podstawowy krok zapewnia, że badanie zachowuje koncentrację i celowość przez cały czas jego trwania.

Krok 2 – Wybór ekspertów: Wybór właściwej grupy ekspertów jest niezbędny dla sukcesu i skuteczności techniki delfickiej. Eksperci ci powinni posiadać odpowiednią wiedzę, umiejętności i doświadczenie dotyczące badanego tematu. Grupa powinna być zróżnicowana, aby zapewnić szerokie spektrum punktów widzenia. Liczba uczestników może się różnić w zależności od wielkości i złożoności badania, ale ogólnie zaleca się uwzględnienie co najmniej 10–15 ekspertów.

Krok 3 – Opracowanie i uruchomienie kwestionariuszy: Ten krok obejmuje opracowanie kwestionariuszy w celu zebrania opinii od ekspertów. Zazwyczaj początkowy kwestionariusz jest otwarty, aby umożliwić ekspertom swobodne dzielenie się swoimi opiniami bez wpływu z zewnątrz. W rundzie 1 eksperci otrzymują kwestionariusz otwarty i niezależnie przekazują spostrzeżenia, przewidywania lub sugestie związane z celami badania. W rundzie 2 moderator podsumowuje i anonimizuje odpowiedzi z rundy 1, aby utworzyć bardziej ukierunkowany kwestionariusz na następną rundę. Jeśli zajdzie taka potrzeba, można przeprowadzić dodatkowe rundy w celu dopracowania opinii na podstawie osiągniętych poziomów konsensusu, kontynuując do momentu osiągnięcia wstępnie zdefiniowanego konsensusu lub podjęcia przez moderatora decyzji o zakończeniu procesu.

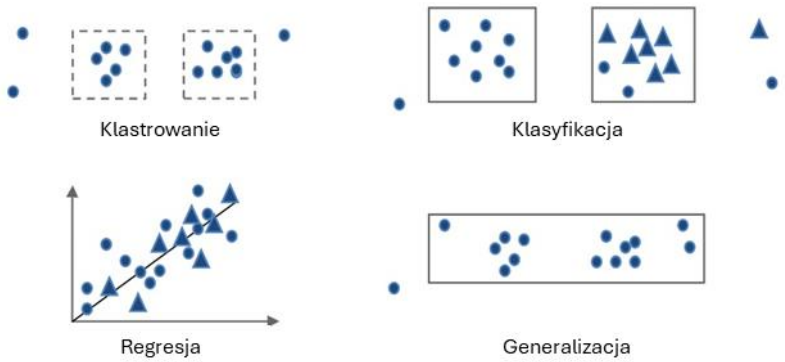
Krok 4 – Wykorzystanie wyniki: Po zakończeniu procesu delfickiego i osiągnięciu konsensusu, wyniki są analizowane i wykorzystywane do podejmowania decyzji, prognozowania, opracowywania polityki organizacji lub innych celów określonych w kroku 1. Anonimowość badań delfickich pomaga zapewnić, że końcowe wyniki są bezstronne i odzwierciedlają łączną wiedzę ekspertów zaangażowanych w badanie.

3.4. Podejście Ilościowej Eksploracji Danych do Odkrywania Wiedzy

Jak pokazano na Rysunku 3.2, ilościowa **Eksploracja Danych jest mocno wspierana przez formalne narzędzia matematyczne, a bardziej bezpośrednio statystyczne**. Możemy tutaj użyć argumentu, że każda operacja statystyczna oparta na danych może być uważana za analizę ilościową. Głównym celem tych operacji jest wyodrębnienie rzeczywistych wzorców z danych i wygenerowanie przydatnych spostrzeżeń w obserwowanym procesie (w przypadku analizy w biznesie lub łańcuchach dostaw). Nie jest to proste zadanie, ponieważ istnieje znacząca niezgodność między założeniami teorii statystycznej/matematycznej a rozkładami i wzorcami, które są obecne w rzeczywistych

danych biznesowych. W praktyce główny przepływ większości analiz biznesowych oparty jest właśnie na tej niezgodności. Stosując metody ilościowe niezwykle ważne jest, aby dane spełniały teoretyczne założenia matematyczne ograniczone przez obserwowany model, tak aby można było traktować wyniki modelu jako ważne i potencjalnie podejmować decyzje na ich podstawie.

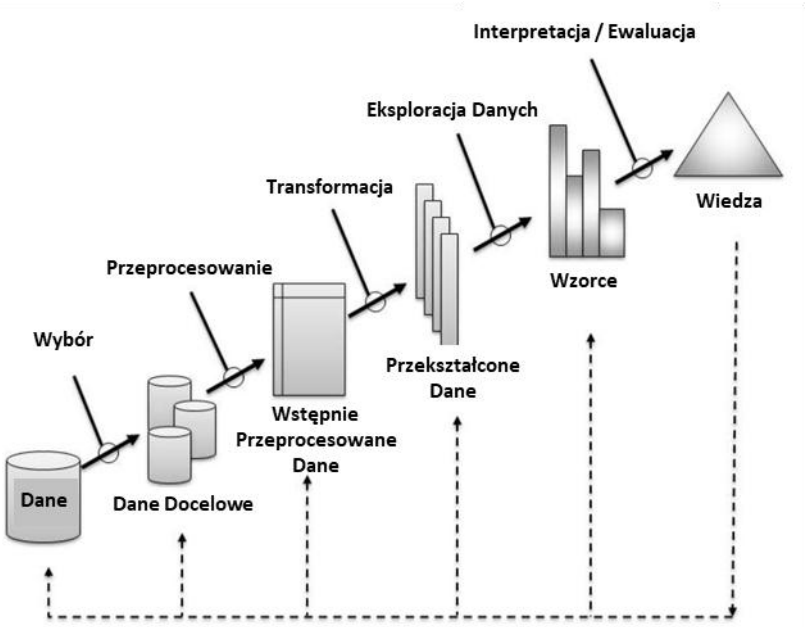
Jak omówiono w poprzedniej sekcji, metody eksploracji danych stanowią podstawowy komponent procesu KDD i są wielokrotnie w nim wykorzystywane. Eksploracja danych jest techniką multidyscyplinarną, której ostateczne metody analityczne opierają się na matematyce. Statystyka odgrywa szczególnie istotną rolę w procesie analizy danych w fazie przygotowywania danych, stanowiąc podstawę dla kilku metod eksploracji danych. Eksploracja danych obejmuje wykorzystanie wydajnych algorytmów w celu odkrycia oczekiwanych lub domniemanych wzorców. Jak pokazano na Rysunku 3.4, zadania Eksploracji Danych można podzielić na pięć kategorii: klastrowanie, klasyfikacja, regresja, reguły asocjacyjne i generalizacja. **Klastrowanie** ma na celu grupowanie obiektów bazy danych tak, aby obiekty w obrębie klastra były do siebie podobne, podczas gdy obiekty w różnych klastrach różniły się od siebie. **Klasyfikacja** obejmuje naukę funkcji, która przypisuje wartości atrybutów do wstępnie zdefiniowanych klas. **Regresja**, metoda statystyczna szacująca relacje między zmiennymi, powszechnie stosowana do przewidywania i prognozowania, znacznie pokrywając się z uczeniem maszynowym. **Reguły asocjacyjne** są stosowane do opisywania silnych relacji w procesach transakcyjnych, takich jak „gdy A i B, to C". **Generalizacja** ma na celu wyrażenie dużej ilości danych w możliwie jak najbardziej zwarty sposób (Su, 2016). Główne techniki stosowane w Data Mining to: reguły klasyfikacji lub drzewa decyzyjne, regresja, klastrowanie, algorytmy genetyczne, modelowanie oparte na agentach itp.



Rysunek 3.4 Zadania Eksploracji Danych

Source: Su (2016)

Wiedzę można wyodrębnić z przetworzonych danych ilościowych. **Proces KDD obejmuje dziewięć kroków** przedstawionych na Rysunku 3.5. Ważne jest, aby zauważyć, że eksploracja danych jest przeprowadzana na bazie przekształconych danych, z których uprzednio wykluczono nieistotne informacje. Wzory odkryte w tym procesie są następnie interpretowane i oceniane w określonym kontekście w celu uzyskania wiedzy, która może wspomóc proces decyzyjny.



Rysunek 3.5 Kroki składające się na proces KDD

Źródło: Fayyad i in. (1996)

Jak już wcześniej wspomniano, powszechnym standardem opisu etapów procesu KDD jest CRISP-DM, który stanowi wiodący model przemysłowy. W odniesieniu do bieżących publikacji i badań przeanalizowanych przez Su (2016), typowe techniki eksploracji danych i zastosowania w łańcuchach dostaw obejmują:

- **Drzewa decyzyjne** (Zastosowanie: rozwiązywanie problemów dostawców, które można sprowadzić np. do zestawu możliwych wyników dla każdej decyzji, wraz z oceną prawdopodobieństwa wystąpienia każdego wyniku).
- **Regresję** (Zastosowanie: prognozowanie i szacowanie popytu klientów na nowy produkt).
- **Regułę asocjacyjną** (Zastosowanie: identyfikowanie przyczyn awarii produktu, optymalizacja zdolności produkcyjnej i umożliwienie utrzymania ruchu opartego na stanie jakościowym).

- **Algorytm genetyczny** (Zastosowanie: rozwijanie hipotezy poprawy działania VMI (Zarządzania Zapasami Przez Dostawcę – ang. Vendor Managed Inventory) w niepewnym środowisku popytu).
- **Algorytmy klastrowania** (Zastosowanie: algorytm centroidów (k-średnich – ang. k-Mean) do kategoryzacji zwracanych towarów w celu poprawy jakości procesów produkcyjnych).
- **System eksploracji danych wieloagentowych** (Zastosowanie: wspieranie decyzji dotyczących planowania produkcji w oparciu o analizę historycznych danych o popycie na produkty).

Wiedza wydobyta z wykorzystaniem Data Mining jest zazwyczaj przechowywana i prezentowana przy użyciu **Systemów Eksperckich (ang. Experts Systems – ES)**. ES to wyrafinowany system wiedzy zaprojektowany w celu naśladowania ludzkiej wiedzy specjalistycznej w różnych obszarach zastosowań. Olson i Courtney (1992) definiują ES jako program komputerowy w określonej domenie, obejmujący pewną ilość sztucznej inteligencji w celu naśladowania ludzkiego myślenia w drodze dojścia do tych samych wniosków, co ekspert danej dziedziny. Komponent ES jest idealny do wspomagania decydenta w obszarze, w którym wymagana jest wiedza specjalistyczna (Turban, 1995). Zasadniczo ES przenosi wiedzę specjalistyczną od eksperta (lub innego źródła) do komputera. Może on albo wspierać decydentów, albo całkowicie ich zastępować i jest najszerzej stosowaną, a także najbardziej udaną komercyjnie technologią sztucznej inteligencji (Turban i in. 2007). Jednym z uzasadnień budowy ES jest dostarczanie wiedzy eksperckiej dużej liczbie użytkowników (Kock, 2005). Według Turbana i in. (2007), ES są uważane za część Systemów Wsparcie Decyzji (ang. Decision Support System – DSS), który można scharakteryzować jako komputerowy system informacyjny łączący modele i dane w celu rozwiązania problemów półstrukturalnych i niestrukturalnych z wysokim poziomem zaangażowania użytkownika (Mirčetić i in. 2016, Turban i in. 2007). Następny rozdział omawia uczenie maszynowe (ang. Machine Learning – ML), które odnosi się do zdolności komputerów do uczenia się i reprezentowania wiedzy pochodzącej z danych wejściowych do procesu. ML można postrzegać jako pomost między wynikami uzyskanymi przez Data Mining a narzędziami Business Intelligence używanymi do prezentowania metryk raportowania dla kadry kierowniczej w formacie umożliwiającym menedżerom podejmowanie świadomych decyzji biznesowych.

BIBLIOGRAFIA

1. Behera, P. C., Dash, C. & Mohapatra, S. (2019). Data Mining and Knowledge Discovery (KDD). International Journal of Research and Analytical Reviews, 6(1), s. 101-106.
2. Fayyad, U. M., Patetsky-Shapiro, G. & Smith, P. (1996). From Data Mining to Knowledge Discovery in Databases. American Association for Artificial Intelligence, 17, s. 37-34.
3. Kock, E. D. (2005) Decentralising the Codification of Rules in a Decision Support Expert Knowledge Base (MSc thesis). University of Pretoria; 2005. Dostępne: http://repository.up.ac.za/handle/2263/22959
4. Lee, P. M. (2013). Use of Data Mining in Business Analytics to Support Business Competitiveness. Review of Business Information Systems, 17(2), s. 53-58.
5. Mirčetić, D., Ralević, N., Nikoličić, S., Maslarić, M. & Stojanović, Đ. (2016). Expert system models for forecasting forklifts engagement in a warehouse loading operation: A case study. Promet-Traffic&Transportation, 28(4), s. 393-401.
6. Naisola-Ruiter, V. (2022). The Delphi technique: a tutorial. Research in Hospitality Management, 12(1), s. 91-97.
7. Olson, D. L., Courtney, J. F. & Courtney, J. F. (1992). Decision support models and expert systems. New York: Macmillan, USA.
8. Paivarinta, T., Pekkola, S. & Moe, C. E. (2011). Grounding Theory from Delphi Studies. In: Proceedings of the 32nd International Conference on Information Systems (ICIS 2011): Research Methods and Philoshophy, 4-7 December 2011, Shanghai, China.
9. Rahman, F. A., Shamsuddin, S. M., Hassan, S. & Haris, N. A. (2016). A Review of KDD-Data Mining Framework and Its Application in Logistics and Transportation. International Journal of Supply Chain Management, 2(1), s. 1-9.
10. Steurer, J. (2011). The Delphi method: an efficient procedure to generate knowledge. Skeletal Radiol, 40, s. 959-961.
11. Su, W. (2016). Knowledge Discovery in Supply Chain Transaction Data by Applying Data Farming (Master thesis). Technical University of Dortmund, Dortmund, DE.
12. Turban, E. (1995). Decision support and expert systems Management support systems. Prentice-Hall, Inc. New York, USA.

13. Turban, E., Rainer, R. K. & Potter, R. E. (2007). Introduction to Information Systems: Supporting and Transforming Business. John Wiley & Sons, Inc. USA.

4. UCZENIE MASZYNOWE (MACHINE LEARNING)

Autor: Dejan Mirčetić

Istnieje wiele pytań na temat tego, czym jest Uczenie Maszynowe (ang. Machine Learning – ML). Czy jest to rzeczywiście proces, w którym maszyny samodzielnie się uczą wykorzystując do tego otoczenie zewnętrzne, czy też jest to sformalizowany za pomocą algorytmów matematycznych proces, który pozwala komputerom „wymyślać" reguły funkcjonujące w świecie zewnętrznym? Jakie narzędzia wykorzystuje ML? Jak wygląda typowy przepływ danych w kanale ML? Czy jest on stosowalny w tradycyjnych branżach, a nie tylko w branżach związanych z IT i Internetem? Jakie miejsce ML zajmuje w kontekście biznesu? Jak systematycznie wykorzystywać je do rozwiązywania problemów biznesowych? Czy istnieje jakaś architektura dotycząca tego, jak stosować uczenie maszynowe w ŁD?

Na te i podobne pytania postaramy się udzielić odpowiedzi w kolejnym rozdziale, kończąc go rzeczywistym przykładem studium przypadku zastosowania algorytmów ML w łańcuchu dostaw żywności.

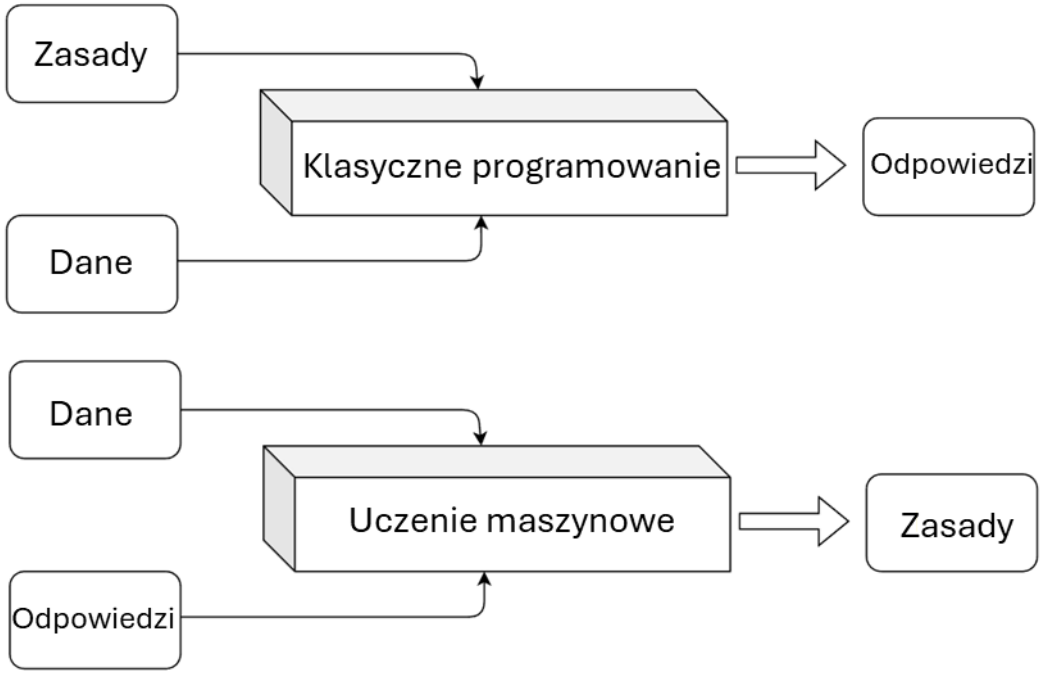
4.1. Czym jest uczenie maszynowe?

Uczenie maszynowe (ML) to dyscyplina skupiająca się na dwóch powiązanych ze sobą pytaniach: Jak można konstruować systemy komputerowe, które automatycznie ulepszają swoje możliwości wykorzystując do tego doświadczenie? oraz Jakie są podstawowe statystyczne prawa teorii informacji obliczeniowej, które rządzą wszystkimi systemami uczącymi się, w tym komputerami, ludźmi i organizacjami? Badanie uczenia maszynowego jest ważne zarówno dla znalezienia odpowiedzi na te podstawowe pytania naukowe i inżynieryjne, jak i dla wysoce praktycznego oprogramowania komputerowego, które zostało wyprodukowane i wdrożone w wielu aplikacjach (Jordan & Mitchell, 2015).

ML wynika z pytania: czy działanie komputera mogłoby wyjść poza „to, co wiemy, jak mu zlecić wykonanie jakiegoś zadania" i nauczyć się samodzielnie, jak wykonać określone czynności? Czy komputer mógłby nas zaskoczyć? Czy komputer mógłby automatycznie nauczyć się tych reguł, opierając się na danych, zamiast o ręcznie napisane

reguły przetwarzania danych stworzone przez programistów? To pytanie otwiera drzwi do nowego paradygmatu programowania (Chollet, 2021).

ML umożliwia fundamentalną zmianę paradygmatu programowania (Rysunek 4.1). W programowaniu klasycznym programista (człowiek) wprowadza reguły (program) oraz dane, które są analizowane i przetwarzane zgodnie z tymi regułami. W rezultacie odpowiedzi są dostarczane na końcu procesu. Z drugiej strony, w ML programista (człowiek) wprowadza dane z odpowiedziami oczekiwanymi od danych, a dopiero na ich podstawie powstają reguły.



Rysunek 4.1 Klasyczne programowanie kontra szkolenie systemów uczenia maszynowego

Źródło: Chollet (2021)

Klasyczne programowanie można rozumieć jako programowanie imperatywne, ponieważ programista wstępnie definiuje wszystkie reguły, a wykonywanie kodu odbywa się zgodnie z nimi, podczas gdy uczenie maszynowe można rozumieć jako programowanie deklaratywne, w którym wyrażamy cele wyższego poziomu lub opisujemy ważne ograniczenia i polegamy na algorytmach matematycznych, aby zdecydować, jak i/lub kiedy przełożyć to na działanie.

Obecnie ML jest podstawą niezliczonych istotnych dla wielu procesów aplikacji, w tym wyszukiwania w sieci, ochrony poczty e-mail przed spamem, rozpoznawania mowy, rekomendacji produktów i innych (Ng, Andrew 2017). Wielu programistów systemów AI zdaje sobie obecnie sprawę, że w przypadku wielu aplikacji o wiele łatwiej jest „szkolić" system, pokazując mu przykłady pożądanego zachowania wejścia-wyjścia, niż programować go

ręcznie, przewidując pożądaną odpowiedź dla wszystkich możliwych danych wejściowych. Efekt wykorzystania ML był również szeroko identyfikowany w całej branży informatycznej i w szeregu branż zajmujących się problemami intensywnie wykorzystującymi dane, takimi jak usługi konsumenckie, diagnostyka usterek w złożonych systemach i kontrola łańcuchów logistycznych (Jordan & Mitchell, 2015).

4.2. Podstawy i założenia teoretyczne ML

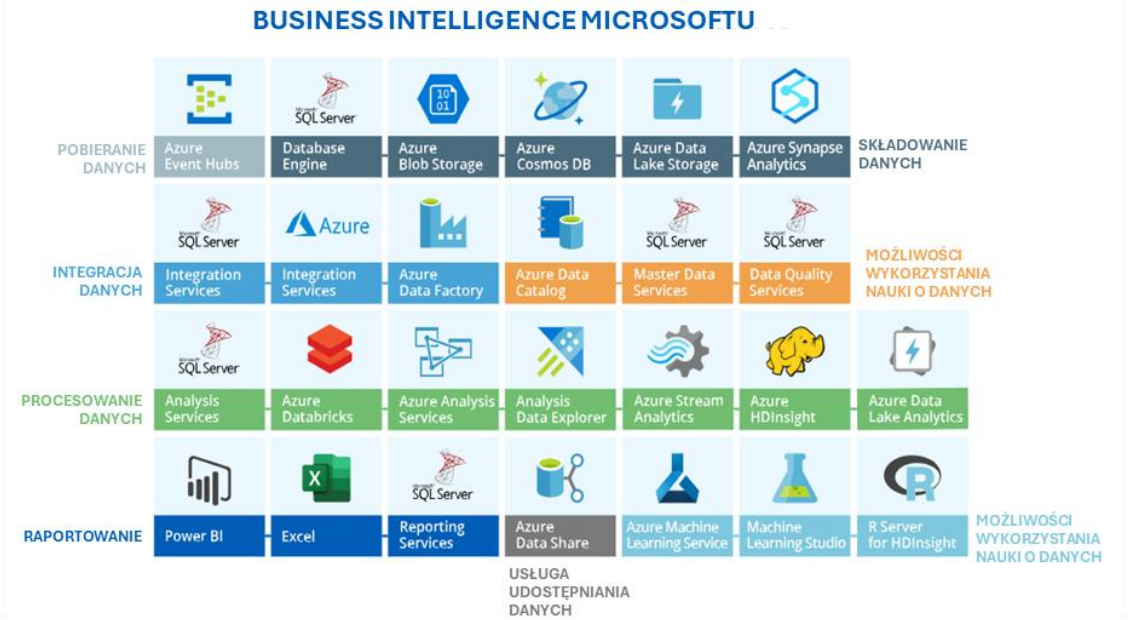
Pochodzenia ML należy szukać w matematyce, a dokładniej w statystyce. Stąd też ML wykorzystuje podstawy teoretyczne i algorytmy opracowane w **uczeniu statystycznym**, jednakże eksperci toczą dyskusję, czy ML jest prawdziwą dziedziną samą w sobie, czy też jest po prostu częścią statystyki. W praktyce algorytmy ML zwykle nie mają pewnego poziomu sztywności matematycznej i czasami łatwo przekraczają pewne ograniczenia matematyczne obecne w statystyce. Na przykład algorytmy ML nie zwracają zbytniej uwagi na przedziały ufności podczas optymalizacji współczynników w algorytmach parametrycznych, chociaż jest to jeden z najważniejszych tematów w statystyce. **Ogólnie rzecz biorąc, ML i statystyka w dużym stopniu się pokrywają**, a niektórzy z najbardziej znanych twórców i profesorów algorytmów ML twierdzą, że jest to po prostu część statystyki (Hastie i in., 2009). Niemniej jednak, będąc dziedziną samą w sobie lub częścią statystyki, ML składa się z kilku kroków w pozyskiwaniu wiedzy z danych. Nie istnieje jednak powszechny konsensus co do tych kroków, ale ogólnie rzecz biorąc, można je przedstawić jako transformację różnych źródeł danych w spostrzeżenia Business Intelligence.

W kontekście biznesowym modele ML są bezużyteczne, bez odpowiedniego wsparcia w zakresie wstępnego przetwarzania danych, eksploracji danych i stosowania spostrzeżeń w rzeczywistych procesach. Dlatego tworzenie algorytmów ML bez możliwości aktualizacji modelu i wykorzystania jego wyników do rzeczywistego procesu podejmowania decyzji nie przynosi żadnej wartości nowoczesnym przedsiębiorstwom. W związku z tym w dzisiejszej analityce biznesowej ilościowy proces ML jest zwykle częścią przepływu pracy Business Intelligence. Dokładniej rzecz biorąc, jest częścią ważnych podprocesów Business Intelligence (nauka o danych i analiza danych stanowią część Business Intelligence). Szczegóły dotyczące roli ML w tych procesach i samych rzeczywistych procesów generowania wartości dla biznesu za pośrednictwem ML zostaną przedstawione w kolejnym podrozdziale.

4.3. Business Intelligence i ML in ŁD

Business Intelligence, w kontekście ŁD, to proces wyciągania wniosków na temat obserwowanych procesów łańcucha dostaw, w oparciu o modelowanie danych generowanych w ramach tych procesów. BI opiera się głównie na statystyce, ale bierze również pod uwagę inne obszary matematyczne: badania operacyjne, algebra liniowa, logika rozmyta (w przypadku, gdy danych jest niewiele lub ich brakuje), optymalizacja numeryczna, metaheurystyka itp. Ponadto nowe technologie przełomowe stają się również ważnym aspektem analizy danych i dostarczania wniosków: **uczenie maszynowe, sztuczna inteligencja, cyfrowe bliźniaki, smartization, żywe laboratoria** itp.

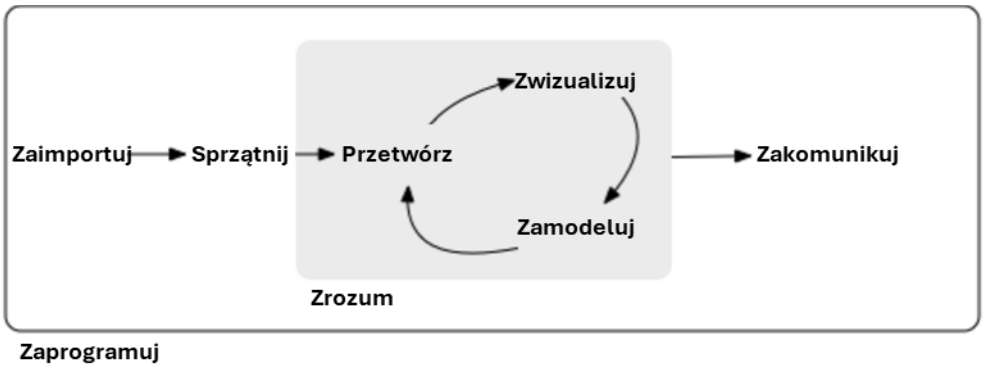
Nie ma ściśle zorganizowanych procedur dotyczących tego, jak powinny być zorganizowane procedury Business Intelligence i przepływy pracy ML, ale istnieją pewne przydatne wytyczne w praktyce i literaturze, które okazały się skuteczne podczas przeprowadzania analizy. Procedura przeprowadzania Business Intelligence jest również zróżnicowana w zależności od źródła oprogramowania, które jest używane do analizy. Na przykład Microsoft oferuje kilka narzędzi za pośrednictwem swojego pakietu Microsoft Business Intelligence, które wykonują różne zadania**:** **pobieranie danych, przechowywanie danych, integrację danych, zarządzanie danymi, przetwarzanie danych, raportowanie, udostępnianie danych oraz naukę o danych** (Rysunek 4.2).



Rysunek 4.2 Architektura Microsoft Business Intelligence

Źródło: ScienceSoft (b.d.)

W danej architekturze procedury ML są stosowane tylko na poziomie nauki o danych za pośrednictwem kilku narzędzi: usług Azure ML, ML studio i R Server dla HDInsight. Ogólna procedura wykonywania analizy danych w kontekście ML w R Server, jest przedstawiona na Rysunku 4.3.

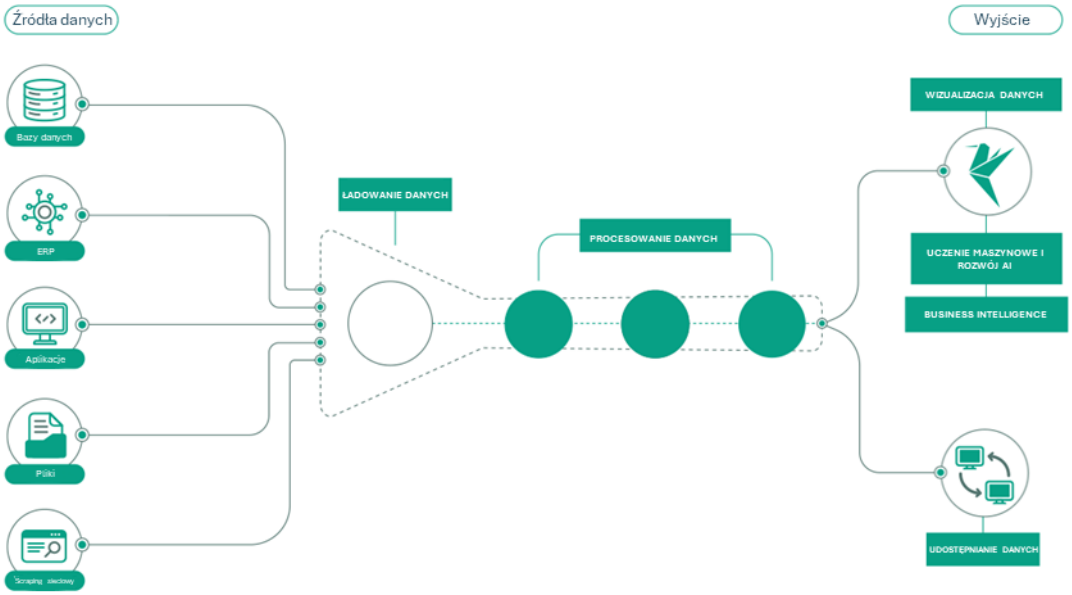


Rysunek 4.3 Etapy analizy danych ML w R

Źródło: Wickham i in. (2023)

W przypadku ML zwykle panuje **błędne przekonanie, że większość czasu i wysiłku poświęca się na faktyczne budowanie algorytmów ML**. W rzeczywistości jest zupełnie odwrotnie, większość czasu poświęca się zazwyczaj na zmagania z danymi i zadaniami wstępnego przetwarzania, a nie na proces modelowania. Czasami wszystkie procesy poprzedzające proces modelowania są o wiele bardziej wymagające i stawiające wyzwania użytkownikom. Dlatego nie ma żadnych przeciwwskazań co do tego, jak te kroki należy wykonać. Rysunek 2.9 przedstawia przykład dobrego podejścia do przekształcania danych w spostrzeżenia biznesowe i wiedzę ogólną. Procedura rozpoczyna się od kroku importu, który jest jednym z najważniejszych kroków w budowaniu modeli ML, ponieważ bez importowania danych do oprogramowania nie można przeprowadzić żadnego rodzaju analizy. Zazwyczaj oznacza to, że należy pobrać dane przechowywane w pliku, bazie danych lub interfejsie programowania aplikacji internetowych (API) i załadować je do ramki danych w R (Wickham i in, 2023). Drugi krok jest związany z uporządkowaniem danych, co jest procedurą unikalną dla R i dotyczy przekształcania danych do określonej formy w celu ich dalszej analizy (każda kolumna jest zmienną, a każdy wiersz jest ramką danych obserwacja-Tibble). Następny krok jest związany z przekształceniem danych, co zwykle obejmuje zawężenie zestawu obserwacji do podpróby zainteresowania. Ponadto może również obejmować tworzenie nowych zmiennych jako kombinacji kilku istniejących zmiennych lub generowanie statystyk podsumowujących. Wizualizacja i modelowanie pełnią odrębne, ale uzupełniające się role

w dziedzinie analizy danych. Wizualizacja jest głęboko zorientowaną na człowieka czynnością, oferującą spostrzeżenia, które mogą umknąć bardziej sformalizowanym podejściom. Dobrze opracowana wizualizacja może ujawnić nieoczekiwane wzorce, skłonić do nowych pytań, a nawet zasugerować, że oryginalne pytania mogą wymagać dopracowania lub innych danych. Natomiast modele zapewniają matematyczne lub obliczeniowe ramy do odpowiadania na precyzyjnie sformułowane pytania. Oferują skalowalność i wydajność, dzięki czemu nadają się do obsługi dużych zestawów danych. Jednakże modele (w których ML jest również uwzględnione) cechują się nieodłącznymi założeniami i nie ma możliwości kwestionowania ani podważania tych założeń. W związku z tym modele mogą nie mieć zdolności do zaskakiwania lub ujawniania nieprzewidzianych spostrzeżeń. Synergia między wizualizacją a modelowaniem jest widoczna w ich wspólnej roli w analizie danych. Wizualizacja pomaga w początkowej eksploracji, zachęcając do formułowania precyzyjnych pytań, podczas gdy modele systematycznie dostarczają odpowiedzi w ramach zdefiniowanych parametrów. Rozpoznanie silnych i słabych stron każdego podejścia jest kluczowe, co prowadzi do bardziej wszechstronnego i świadomego procesu analizy danych. Ostatni krok stanowi komunikację, która jest niezbędna do osiągniecia sukcesu analizy danych, ponieważ jeśli informacje nie zostaną dostarczone decydentowi w odpowiedni i spójny sposób, cała analiza może pójść na marne. Kluczowym elementem w analityce danych są modele ML, bez których nie można by wyciągnąć wniosków na temat procesów biznesowych. Aby rozwiązać konkretne problemy ŁD, architektura ogólnych aplikacji Business Intelligence (przedstawiona na Rysunku 4.2) musi zostać lepiej dostosowana, podobnie jak modele ML. W związku z tym, aby przekształcić sposób działania łańcuchów dostaw poprzez zwiększenie efektywności operacyjnej, usprawnienie podejmowania decyzji i dążenie do realizacji celów korporacyjnych, organizacja **Equilibrium AI** opracowała platformę AI i ML przedstawioną na Rysunku 4.4.



Rysunek 4.4 Proces przetwarzania danych i wiedzy ML dla firmy Equilibrium AI

Źródło: Equilibrium AI (b.d.)

Rysunek przedstawia dobry przykład codziennej praktyki, w jaki sposób ekstrakcja wiedzy i spostrzeżeń jest generowana w aplikacjach ŁD. Zasadniczo proces składa się z operacji back-end i front-end w celu tworzenia wartości (spostrzeżeń biznesowych) dla użytkowników. Proces back-end rozpoczyna się od ekstrakcji danych z różnych źródeł danych, które zwykle znajdują się w ŁD:

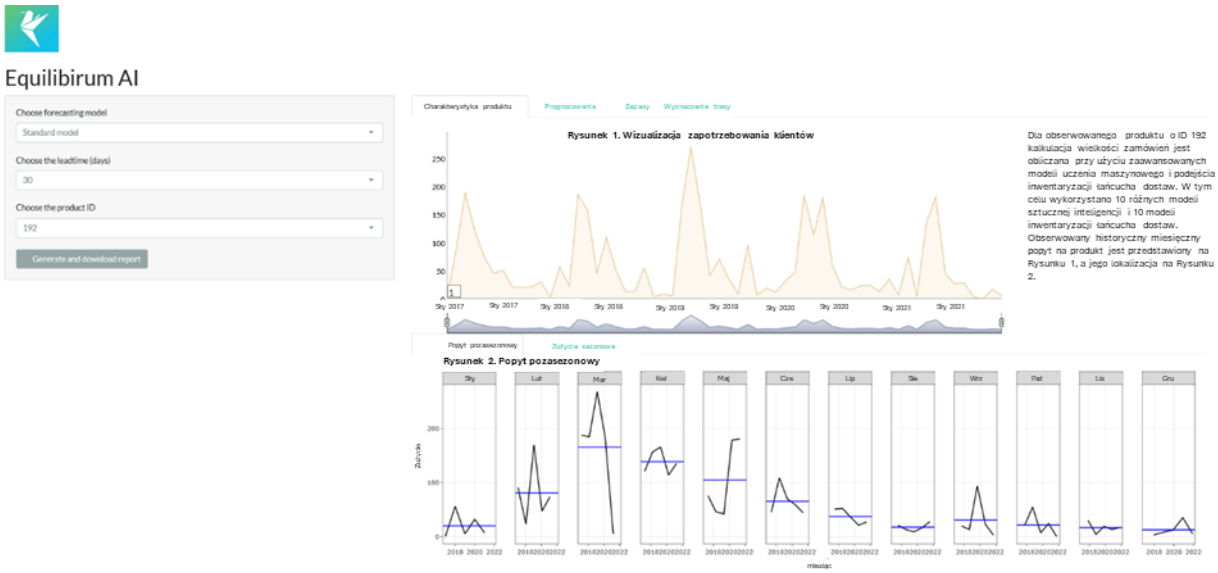
- Bazy danych;
- Systemy planowania zasobów przedsiębiorstwa (SAP, Navigator, Microsoft Dynamics itp.);
- Aplikacje (internetowe API);
- Pliki płaskie (csv, xlsx, JSON itp.);
- Sieć, Internet i inne źródła online.

Każde ze źródeł danych ma inną strukturę, protokoły i odpowiednio procedury dotyczące sposobu ekstrakcji i ładowania danych w celu oczyszczenia i wstępnego przetworzenia przed zastosowaniem algorytmów ML. W związku z tym proces ten jest wykonywany za pomocą narzędzi ładujących dane, które posiadają wstępnie zaprogramowany kod służący eksploracji danych pochodzących z różnych źródeł i przekazywania danych wierszowych do nowej bazy danych, która jest ustrukturyzowana

i zorganizowana w celu możliwości zastosowania modeli ML. Przed zastosowaniem modeli ML istnieje jeden dodatkowy krok zwany wstępnym przetwarzaniem danych. W tym kroku dane wierszowe zebrane od firm są sprawdzane pod kątem błędnych danych wejściowych, wartości nielogicznych, prawidłowej struktury danych wejściowych, obserwacji odstających, powtórzeń, NA, NaN itp. Procedura jest kontynuowana poprzez scalanie danych zewnętrznych z danymi firmy. Dane te są zazwyczaj powiązane z czynnikami zewnętrznymi, które mogą potencjalnie wpływać na obserwowany proces biznesowy ŁD, na przykład dane pogodowe, wskaźnik cen konsumpcyjnych, średni dochód w danym regionie, specyficzne cechy demograficzne w danym obszarze, ceny gazu, fale pandemii, komentarze w mediach społecznościowych na temat produktów firmy itp. Jest to bardzo istotne z uwagi na holistyczne zbieranie wszystkich możliwych czynników (wewnętrznych i zewnętrznych), które mogą wpływać na dany proces biznesowy, co zwiększa szansę, że modele ML znajdą właściwy sygnał w danych i będą w stanie wyciągnąć właściwe wnioski i reguły jakie są główne przyczyny, dla których proces biznesowy zachowuje się tak jak zidentyfikowano w ramach obserwacji.

Po połączeniu danych wewnętrznych i zewnętrznych, wstępne przetwarzanie obejmuje wykrywanie sygnału, usuwanie „szumów" z danych, inżynierię cech i losowe dzielenie danych na trening i test (czasem na dane walidacyjne, jeśli opracowano model sieci neuronowej). Dane, które wychodzą z etapu wstępnego przetwarzania, są czyszczone i strukturyzowane w celu możliwości zastosowania modeli ML.

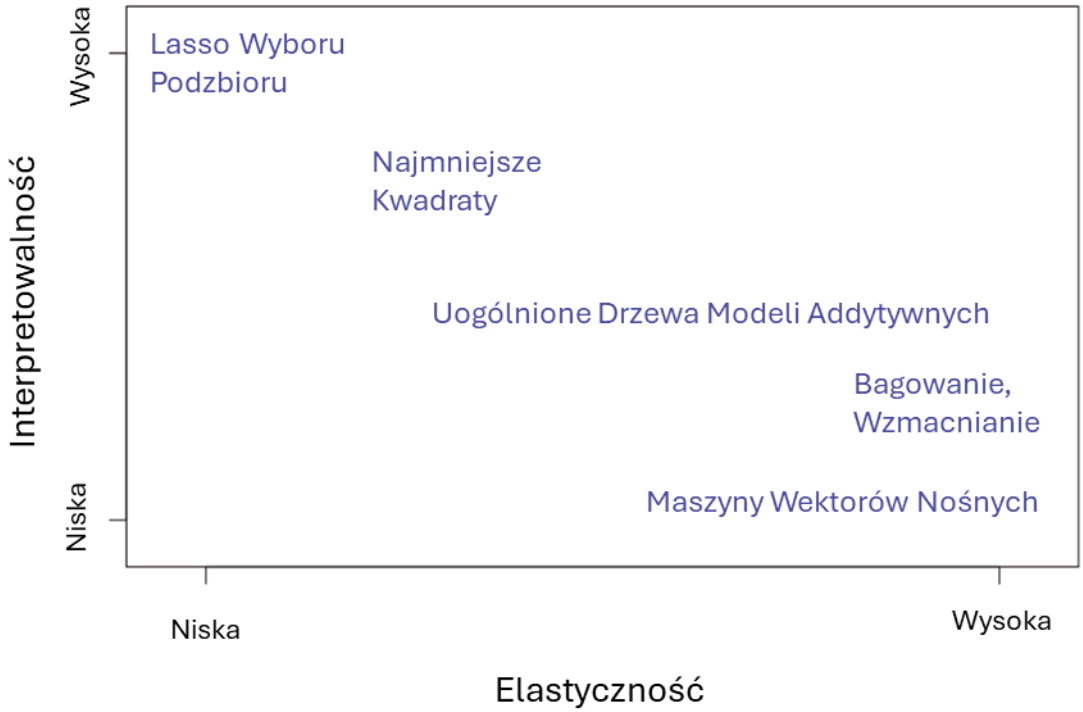
Część wyjściowa składa się z wizualizacji danych, rozwoju ML i AI oraz udostępniania danych. Czasami dane te są udostępniane bez stosowania modeli ML na innych platformach, które przeprowadzają różnego rodzaju analizy (tylko raportowanie do interesariuszy lub agencji rządowych). Proces wizualizacji jest wykonywany za pośrednictwem części front-end platformy, która jest zorientowana na użytkownika i pozwala użytkownikom wykonywać zapytania dotyczące tego, jakie dane łańcucha dostaw, w jaki sposób i w jakich ustawieniach chcą zobaczyć (na przykład Rysunek 4.5).



Rysunek 4.5 Typowa część wizualizacyjna platformy ML w ŁD

Źródło: Opracowanie własne

Z drugiej strony, część przetwarzania danych ML nie jest widoczna dla użytkowników i jest ona niełatwa do zrozumienia. Dlatego modele ML są czasami uważane za **modele czarnej skrzynki**, w ramach których nie ma jasnego zrozumienia, w jaki sposób dokładnie narzędzie połączyło obserwowane dane wejściowe z obserwowanymi danymi wyjściowymi. Jest to jedna z przeszkód, która uniemożliwia szersze wykorzystanie modeli ML w praktyce, zwłaszcza tych, które cechują się dużym stopniem złożoności interpretacji (Rostami-Tabar & Mircetic, 2023). W związku z tym możemy podzielić modele ML na te o wysokiej interpretowalności – niskiej elastyczności i niskiej interpretowalności - wyższej elastyczności (Rysunek 4.6). Ogólnie mówiąc, wraz ze wzrostem elastyczności metody ML, zwykle wzrasta dokładność modelu ML, a jego interpretowalność maleje (Mirčetić i in., 2016).



Rysunek 4.6 Reprezentacja kompromisu (tradeoff) między elastycznością a interpretowalnością przy użyciu różnych metod uczenia maszynowego

Źródło: Hastie i in. (2009)

Dane biznesowe ML i ŁD

Jeśli użytkownicy lepiej rozumieją wizualizację i grafikę, taką jak na Rysunku 4.5, to jaki jest cel wykorzystania ML i czy możliwe jest pominięcie modelowania danych za pomocą algorytmów uczenia maszynowego i zastąpienie ich grafikami informacyjnymi? Niestety nie. Być może głównym powodem, dla którego modele ML są potrzebne, jest fakt, iż nie we wszystkich sytuacjach możliwe jest uzyskanie łatwo czytelnych i wykrywalnych wzorców w danych prezentowanych za pomocą grafiki (jak na Rysunku 4.5). Bardziej powszechną sytuacją jest taka, w której grafika zazwyczaj nie może ujawnić tajemnicy tego, co dzieje się w obserwowanych danych biznesowych ŁD. Dlatego tez potrzebne jest wykorzystanie silniejszych narzędzi w postaci algorytmów ML, celem głębszego wnikania w dane i wyszukiwania **reguł generowania danych** (Rysunek 4.7).

Rysunek 4.7 Charakterystyka statystyczna produktów w łańcuchu dostaw żywności (podsumowanie dla wszystkich produktów)

Źródło: Opracowanie własne

Wyciągniecie prostych wniosków z Rysunku 4.7 i wyprowadzenie reguł biznesowych dotyczących procesu generowania danych jest bardzo trudne. Aby zidentyfikować wzorce zaszyte w danych pochodzących z rysunku, opracowany został model ML, który można wykorzystać do podsumowania cech i wykrycia ważnych sygnałów ukrytych w danych. W związku z tym w Równaniu 1 przedstawiono opracowany model ML dla łańcucha dostaw żywności. Podstawowym czynnikiem napędowym i kręgosłupem tego modelu ML jest autoregresyjny zintegrowany model średniej ruchomej, którego ogólna postać jest następująca:

$$y_t' = c + (\phi_1 y_{t-1}' + \ldots + \phi_p y_{t-p}') + (\theta_1 e_{t-1} + \ldots + \theta_q e_{t-q}) + e_t; \quad (1)$$

$$y_t - y_{t-1} = c + \phi_1 (y_{t-1} - y_{t-2}) + \ldots + \phi_p (y_{t-p} - y_{t-p-1}) + (\theta_1 B e_t + \ldots + \underbrace{\theta_q e_{t-q}}_{\theta_q B^q e_t} + e_t);$$

$$y_t - B y_t = c + \phi_1 (y_{t-1} - B y_{t-1}) + \ldots + \phi_p (y_{t-p} - B y_{t-p}) + (e_t (1 + \theta_1 B + \ldots + \theta_q B^q));$$

$$(1 - B) y_t = c + \phi_1 (1 - B)(y_{t-1}) + \ldots + \phi_p (1 - B) y_{t-p} + e_t (1 + \theta_1 B + \ldots + \theta_q B^q);$$

$$(1 - B) y_t = c + \phi_1 (1 - B) B y_t + \ldots + \phi_p (1 - B) B^p y_t + e_t (1 + \theta_1 B + \ldots + \theta_q B^q);$$

$$\underbrace{(1 - B)^d y_t}_{\substack{differencing \\ d\_\deg ree}} \cdot \underbrace{(1 - \phi_1 B - \ldots - \phi_p B^p)}_{AR(p)} = c + \underbrace{e_t (1 + \theta_1 B + \ldots + \theta_q B^q)}_{MA(q)} \cdot$$

Model ML wyraźnie pokazuje swoją niską interpretowalność i cechy czarnej skrzynki. Przeciętnemu użytkownikowi biznesowemu trudno jest zrozumieć powiązania między danymi wejściowymi i wyjściowymi. Co więcej, przeciętny użytkownik biznesowy po zapoznaniu z przedstawionym modelem powinien zadać pytanie: Czym jest równanie (1)? Można argumentować, że równanie (1) przedstawia reguły umieszczone na Rysunku 4.1, wygenerowane przez dane ML i potok wiedzy, które ujawniają tajemnicę procesów generowania danych w danym środowisku biznesowym ŁD.

Na pierwszy rzut oka opracowany model ML w równaniu (1) nie wydaje się przyczynić do poprawy zrozumienia danych. Nadal występuje pewna doza dezorientacji, jak w przypadku Rysunku 4.7, jednakże model ML ma kluczową przewagę nad rysunkiem. W istocie model ML jest **wzorem matematycznym**, który może nie być łatwo zrozumiany przez użytkownika, ale jest całkowicie zrozumiały dla narzędzia komputerowego, które **można zaprogramować tak, aby używało dany wzór** i podejmowało **decyzje biznesowe** w oparciu o odkryte reguły.

BIBLIOGRAFIA

1. Chollet, F. (2021). Deep learning with Python. Simon and Schuster.
2. Equilibrium AI (n.d.). Equilibrium AI Data Pipeline [dostępne: https://eqains.com/, dostęp: 23 stycznia 2024]
3. Hastie, T., Tibshirani, R., Friedman, J. H., & Friedman, J. H. (2009). The elements of statistical learning: Data mining, inference, and prediction (Vol. 2). Springer.
4. Jordan, M. I. & Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. Science, 349(6245), s. 255-260.
5. Mirčetić, D., Ralević, N., Nikoličić, S., Maslarić, M. & Stojanović, Đ. (2016). Expert system models for forecasting forklifts engagement in a warehouse loading operation: A case study. Promet-Traffic & Transportation, 28(4), s. 393-401.
6. Ng, A. (2017). Machine learning yearning. [dostępne: http://www.mlyearning.org/, dostęp: 23 stycznia 2024]
7. Rostami-Tabar, B. & Mircetic, D. (2023). Exploring the association between time series features and forecasting by temporal aggregation using machine learning. Neurocomputing, 548, 126376.
8. ScienceSoft (b.d.). Microsoft Business Intelligence to Drive Robust Analytics and Insightful Reporting dostępne: https://www.scnsoft.com/services/business-intelligence/microsoft, dostęp: 23 stycznia 2024]
9. Wickham, H., Çetinkaya-Rundel, M. & Grolemund, G. (2023). R for data science. O'Reilly Media, Inc.

5. ZARZĄDZANIE PROCESAMI BIZNESOWYMI I EKSPLORACJA PROCESÓW

Autor: Dario Šebalj

Aby pozostać konkurencyjnym w dzisiejszym środowisku biznesowym, kluczowe znaczenie ma skuteczne zarządzanie i stałe doskonalenie procesów. W tym rozdziale omówiono pojęcia Zarządzania Procesami Biznesowymi (BPM – ang. Business Process Management) oraz Eksploracja Procesów (ang. Process Mining), dwie ważne części Business Intelligence, które pomagają firmom analizować, optymalizować i ulepszać swoje procesy operacyjne.

BPM zapewnia zorganizowaną i ustrukturyzowaną metodę identyfikacji, projektowania, wykonywania, monitorowania i ulepszania procesów biznesowych, jednocześnie dostosowując je do celów strategicznych organizacji. Z drugiej strony eksploracja procesów jest narzędziem do identyfikowania i ulepszania rzeczywistych procesów poprzez ekstrakcję wiedzy z dzienników zdarzeń znajdujących się w nowoczesnych systemach informacji biznesowej. Połączenie zarządzania procesami biznesowymi i eksploracji procesów umożliwia obiektywną, opartą na danych metodę zrozumienia i ulepszania procesów biznesowych.

Wykorzystując te metodologie, organizacje mogą znaleźć ukryte nieefektywności i problemy, dostosować się do zmieniających się wymagań rynku i poprawić swoją wydajność oraz zadowolenie klientów. Podstawowe koncepcje, metodologie, narzędzia i rzeczywiste zastosowania BPM i eksploracji procesów zostaną omówione w tym rozdziale.

5.1. Proces biznesowy

Każda organizacja, niezależnie od wielkości lub sektora, jest złożonym systemem powiązanych ze sobą procesów. Procesy te to ustrukturyzowane działania podejmowane w celu osiągnięcia określonego celu organizacyjnego. Na przykład w przedsiębiorstwie produkcyjnym kluczowe procesy mogą obejmować projektowanie produktu, zaopatrzenie w surowce, produkcję, kontrolę jakości i dystrybucję. W przedsiębiorstwie zorientowanym na usługi, takim jak bank, procesy obejmują otwieranie kont, przetwarzanie pożyczek, obsługę

klienta i kontrolę zgodności. Organizacje korzystają z procesów na co dzień, a procesy te mogą być tak różnorodne, jak same organizacje. W szpitalu procesy obejmują przyjęcie pacjenta, leczenie i wypisanie ze szpitala. W placówce edukacyjnej obejmują one rejestrację studentów, prowadzenie kursów i administrację egzaminów. Każdy proces to sekwencja kroków, obejmująca różne działy i personel, często wspierana przez technologię.

Według Dumasa i in. (2018) każdy proces biznesowy składa się z kilku zdarzeń i działań. **Zdarzenia** odpowiadają elementom, które nie mają czasu trwania i dzieją się automatycznie (np. „Otrzymano zamówienie"). Z drugiej strony **czynności** to zadania lub operacje, które są ze sobą powiązane i których wykonanie spełnia cel procesu biznesowego (np. „Zapłać fakturę"). Typowy proces, oprócz zdarzeń i czynności, obejmuje **decyzje**, które wskazują etap, na którym proces decyduje, w którym kierunku pójdzie w przyszłości. Na przykład w procesie sprzedaży jednym z punktów decyzyjnych może być moment, w którym sprzedawca sprawdza, czy produkt jest dostępny w magazynie. Jeśli produkt jest dostępny w magazynie, proces przechodzi do następnej czynności. Jeśli nie ma produktu w magazynie, proces przebiega w inny sposób (np. informując klienta, że zamówienie nie może zostać zrealizowane). Ważnymi częściami procesu są aktorzy/uczestnicy i obiekty. **Aktorzy** to osoby, organizacje lub systemy oprogramowania, które wykonują czynności procesowe, natomiast **obiekty** to urządzenia, materiały, dokumenty papierowe (obiekty fizyczne), dokumenty elektroniczne i zapisy (obiekty informacyjne).

Dumas i in. (2018) twierdzą, że wykonanie procesu skutkuje jednym lub większą liczbą **wyników**. Wynik powinien teoretycznie przynosić korzyści wszystkim stronom zaangażowanym w proces (*pozytywny wynik*). Czasami wartość ta jest osiągana tylko częściowo lub nigdy (*negatywny wynik*).

Von Scheel i in. (2015) definiują **proces biznesowy** jako „zbiór zadań i czynności (operacji i czynności biznesowych) składający się z pracowników, materiałów, maszyn, systemów i metod, które są ustrukturyzowane w taki sposób, aby zaprojektować, stworzyć i dostarczyć produkt lub usługę konsumentowi".

Zrozumienie procesu to dopiero początek. Prawdziwym problemem i szansą jest zarządzanie tymi procesami w sposób systematyczny i zaplanowany. To prowadzi nas do następnego rozdziału: Zarządzanie procesami biznesowymi (BPM). W tej sekcji przyjrzymy się podejściom i ramom, które pozwalają organizacjom nie tylko zarządzać, ale także wykonywać swoje procesy. BPM to coś więcej niż tylko rejestrowanie i analiza procesów; to

kompleksowa metoda opracowywania, wdrażania, monitorowania i ciągłej poprawy procesów biznesowych.

5.2. Zarządzanie Procesami Biznesowymi

W literaturze naukowej i zawodowej możemy znaleźć różne definicje Business Process Management. Gartner (b.d.) definiuje BPM jako „dyscyplinę, która wykorzystuje różne metody do odkrywania, modelowania, analizowania, mierzenia, ulepszania i optymalizacji procesów biznesowych". Według Camundy (b.d.) BPM to „systemowe podejście do rejestrowania, projektowania, wykonywania, dokumentowania, mierzenia, monitorowania i kontrolowania zarówno zautomatyzowanych, jak i niezautomatyzowanych procesów w celu spełnienia celów i strategii biznesowych firmy". Swenson i Rosing (2015) zaproponowali szerszą i być może najbardziej precyzyjną definicję: „Zarządzanie procesami biznesowymi (BPM) to dyscyplina obejmująca dowolną kombinację modelowania, automatyzacji, wykonywania, kontroli, pomiaru i optymalizacji przepływów działań biznesowych w odpowiedniej kombinacji w celu wspierania celów przedsiębiorstwa, obejmująca granice organizacyjne i systemowe oraz angażująca pracowników, klientów i partnerów w obrębie przedsiębiorstwa i poza jego granicami".

Według Freunda i Rückera (2012) nowe projekty BPM często obejmują jeden z tych scenariuszy:

1. Usprawnianie procesów przy użyciu technologii informatycznych (IT).
2. Dokumentacja bieżących procesów.
3. Wprowadzenie całkowicie nowych procesów.

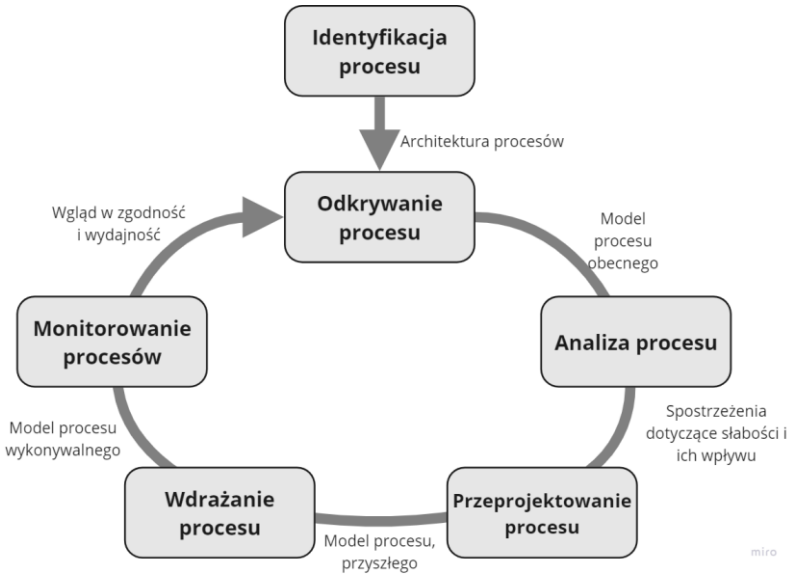
Dumas i in. (2018) see BPM as a continuous cycle comprising the following phases:

- **Identyfikacja procesu** – W tym kroku określany jest problem biznesowy. Procesy, które są ważne dla rozwiązywanego problemu, są identyfikowane, definiowane i łączone. Wynikiem identyfikacji procesu jest nowa lub ulepszona architektura procesu. Ta architektura pokazuje wszystkie procesy organizacji i sposób, w jaki są one ze sobą połączone. Służy do wybierania, który proces lub zestaw procesów ma być obsługiwany przez resztę cyklu życia.
- **Odkrywanie procesów (modelowanie procesów „takich, jakie są – ang. As- is")** – Polega na dokumentowaniu bieżącego stanu wszystkich ważnych

procesów, zwykle w formie jednego lub większej liczby modeli procesów „takich, jakie są".

- **Analiza procesu** – Podczas tego kroku problemy z obecnym procesem As-is są identyfikowane, dokumentowane i, jeśli to możliwe, mierzone za pomocą wskaźników wydajności. Wynikiem tego kroku jest ustrukturyzowana lista problemów. Problemy te są klasyfikowane według możliwego wpływu i szacowanego wysiłku potrzebnego do ich rozwiązania.
- **Przeprojektowanie procesu (modelowanie procesu przyszłego – ang. To-be)** – Celem tej fazy jest znalezienie modyfikacji procesu, które umożliwią przedsiębiorstwu osiągnięcie celów wydajnościowych, a jednocześnie zajęcie się problemami zidentyfikowanymi w poprzedniej fazie. Ta faza zwykle skutkuje modelem procesu To-be.
- **Wdrożenie procesu** – Dostosowania niezbędne do przeniesienia procesu As-is do procesu To-be są planowane i przeprowadzane w tej fazie. Automatyzacja i zarządzanie zmianą organizacyjną to dwa aspekty wdrożenia procesu. Termin „zarządzanie zmianą organizacyjną" opisuje zbiór działań niezbędnych do zmiany sposobu pracy wszystkich uczestników zaangażowanych w proces. Tworzenie i wdrażanie systemów informatycznych (lub ulepszonych wersji obecnych systemów informatycznych) w celu wsparcia przyszłego procesu jest określane jako automatyzacja procesów.
- **Monitorowanie procesu** – po wdrożeniu przeprojektowanego procesu zbierane są a następnie analizowane odpowiednie dane w celu oceny wydajności procesu. Działania korygujące są inicjowane po zidentyfikowaniu wąskich gardeł, powtarzających się błędów lub odchyleń od zamierzonego zachowania.

Cykl opisany powyżej musi być powtarzany w sposób ciągły, ponieważ w bieżącym procesie lub innych procesach mogą pojawić się nowe problemy. Cykl życia BPM pokazano na Rysunku 5.1.



Rysunek 5.1 Cykl życia BPM

Źródło: Dumas i in. (2018)

Freund i Rücker (2012) wymieniają kilka ról, które są zaangażowane w projekty BPM:

- **Właściciel procesu** – osoba, która ma strategiczną odpowiedzialność za procesy. Posiada uprawnienia budżetowe i często jest członkiem pierwszego lub drugiego szczebla zarządzania. Na przykład właścicielem procesu może być CEO firmy.
- **Menedżer procesów** – osoba, która ponosi odpowiedzialność operacyjną za procesy. Często jest menedżerem niskiego lub średniego szczebla. Na przykład menedżer sprzedaży może być menedżerem procesów.
- **Uczestnik procesu** – osoba pracująca z procesem i tworząca wartość (np. sprzedawca).
- **Analityk procesów** – osoba rozumiejąca w sposób ogólny BPM i BPMN w szczegółach, będąca centrum każdego projektu BPM.

BPM pomaga przedsiębiorstwom dopasować swoje procesy do ich ogólnych celów, stać się bardziej wydajnymi i dostosować się do zmieniających się środowisk. W następnej sekcji zostaną przedstawione metody i narzędzia, które są używane do tworzenia dokładnych modeli procesów biznesowych.

Modelowanie procesów biznesowych ma na celu nie tylko rysowanie diagramów, ale także uchwycenie podstawowych procesów w sposób, który ułatwia ich zrozumienie, komunikację i analizę. Interesariusze mogą go używać do wizualizacji złożonych procesów, dostrzegania nieefektywności i wąskich gardeł oraz konceptualizacji usprawnień i innowacji.

W następnej sekcji zostanie zaprezentowana najpopularniejsza metoda modelowania BPMN (Notacja i Model Procesu Biznesowego – ang. Business Process Model and Notation). Omówione zostanie, w jaki sposób to narzędzie może być wykorzystane do efektywnego dokumentowania procesów biznesowych.

5.3. Modelowanie Procesów Biznesowych

Aby zapewnić standaryzowaną, graficzną notację do dokumentowania, projektowania i analizowania procesów biznesowych, wprowadzono **Notację i Model Procesu Biznesowego (BPMN – Business Process Model and Notation)**. Według Lucidcharta (b.d.), Business Process Management Initiative (BPMI) stworzyło Notację Modelowania Procesów Biznesowych (ang. Business Process Modeling Notation), którą poddano wielu zmianom. Inicjatywę przejęła Object Management Group (OMG) po połączeniu się z BPMI w 2005 r. OMG wydało BPMN 2.0 i zmieniło nazwę metody na Notację i Model Procesu Biznesowego (Business Process Model and Notation). Dzięki szerszemu zakresowi symboli i notacji dla diagramów procesów biznesowych działanie to ustanowiło bardziej kompleksowy standard modelowania procesów biznesowych.

BPMN reprezentuje cztery kategorie elementów (Lucidchart, b.d.; Freund i Rücker, 2012):

- **Obiekty przepływu**: zdarzenia, zadania (czynności) i bramki;
- **Obiekty łączące**: przepływ sekwencji, przepływ wiadomości i połączenia;
- **Uczestnicy**: baseny i tory;
- **Artefakty**: obiekty danych, magazyn danych i adnotacje.

5.3.1. Zdarzenia

Aagesen i Krogstie (2015) definiują zdarzenia jako coś, co dzieje się w procesie. W BPMN istnieją trzy typy zdarzeń: początkowe, pośrednie i końcowe. Zdarzenie początkowe jest wyzwalaczem rozpoczęcia procesu. Zdarzenia pośrednie występują w trakcie procesu biznesowego i często oznaczają kamienie milowe lub oczekiwania w procesie. Zdarzenia końcowe oznaczają koniec procesu biznesowego. Są one reprezentowane przez okręgi.

Rysunek 5.2 Oznaczenia zdarzeń początkowych, pośrednich i końcowych

Źródło: Opracowanie własne

Według Dumasa i in. (2018) zdarzenie powinno być nazwane jako [obiekt] + [czasownik w imiesłowach czasu przeszłego]. Oto kilka przykładów, jak nazwać zdarzenia: „Faktura wysłana", „Zamówienie złożone", „Produkty otrzymane"".

Tabela 5.1. przedstawia różne typy zdarzeń początkowych, pośrednich i końcowych. (OMG, 2006).

Tabela 5.1 Typy zdarzeń

Typ	Opis	Symbol
Zdarzenie początkowe		
Brak typu	Rodzaj zdarzenia nie jest wyświetlany.	
Komunikat	Otrzymuje wiadomość od uczestnika, która uruchamia proces.	
Zegarowe	Proces jest uruchamiany o określonej porze (np. w każdy poniedziałek o godz. 9:00).	
Warunkowe	Zdarzenie jest wyzwalane, gdy spełniony jest pewien warunek (np. gdy poziom zapasów jest niższy niż 500 sztuk).	
Zdarzenie pośrednie		
Brak typu	Rodzaj zdarzenia nie jest wyświetlany.	
Komunikat	Wiadomość przychodzi od uczestnika i wyzwala zdarzenie. Powoduje to kontynuację procesu, jeśli oczekiwał na wiadomość.	
Zegarowe	Może działać jako mechanizm opóźniający. Na przykład, jeśli proces oczekuje na dostawę produktu.	
Zdarzenie końcowe		
Brak typu	Rodzaj zdarzenia nie jest wyświetlany.	
Komunikat	Na koniec procesu uczestnikowi wysyłana jest wiadomość.	
Błąd	Na końcu procesu powinien zostać wygenerowany błąd.	
Zerwanie	Wszystkie czynności w procesie powinny zostać natychmiast zakończone.	

Źródło: OMG (2006)

5.3.2. Zadanie (czynność)

Czynność to coś, co jest wykonywane podczas procesu, czynności wykonywane przez osobę lub system. Jest reprezentowane przez prostokąt z zaokrąglonymi rogami.

W BPMN istnieje specjalny podzbiór regularnych zadań zwany podprocesem. Jest on reprezentowany przez prostokąt ze znakiem „+" na dole. Służy on do reprezentowania procesu w procesie. W ten sposób złożoność głównego procesu, tj. procesu będącego w centrum uwagi, jest zmniejszona.

Rysunek 5.3 Oznaczenia czynności i podprocesów

Źródło: Opracowanie własne

Czynność powinna zostać nazwana jako [czasownik w trybie rozkazującym] + [obiekt] (Dumas i in., 2018). Oto kilka przykładowych zadań: „Wyślij fakturę" lub „Prześlij zamówienie".

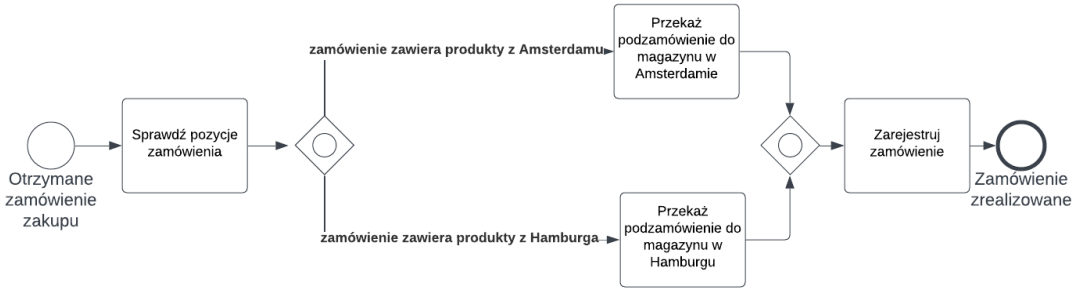
5.3.3. Bramki

Bramki to miejsca, w których procesy dzielą się lub łączą. Są one reprezentowane przez kształt rombu. Istnieją trzy najczęstsze typy bramek: bramka ALBO (wykluczająca), bramka LUB (separacji) i bramka ORAZ (równoległa).

Rysunek 5.4 Notacje bramek LUB, ALBO, ORAZ

Źródło: Opracowanie własne

Według von Rosing i in. (2015), **bramka LUB**, podczas dzielenia, pozwala na aktywację jednej lub więcej gałęzi, w zależności od warunków. Przed scaleniem wszystkie aktywne gałęzie przychodzące muszą zostać ukończone, aby móc kontynuować przepływ. Przykład bramki LUB pokazano na Rysunku 5.5.



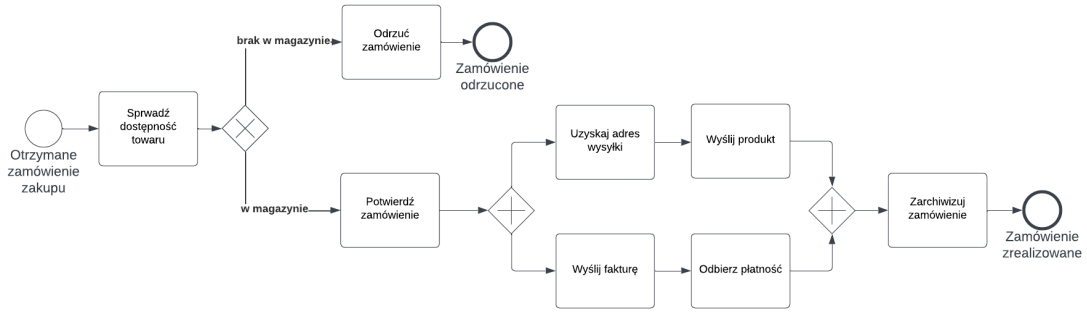
Rysunek 5.5 Przykład użycia bramki LUB

Źródło: Dumas i in. (2018)

W tym przykładzie przedsiębiorstwo posiada magazyny w Amsterdamie i Hamburgu, w których przechowuje różne produkty. Po otrzymaniu zamówienia jest ono dzielone między te magazyny: podzamówienie jest wysyłane do Amsterdamu, jeśli przechowywane są tam określone produkty, a także podzamówienie jest wysyłane do Hamburga, jeśli przechowywane są tam określone produkty. Następnie procedura kończy się, gdy zamówienie zostanie zarejestrowane (Dumas i in., 2018). Możemy zobaczyć, że proces może przebiegać w obu kierunkach (jeśli zamówione produkty są przechowywane w obu magazynach) lub tylko w jednym kierunku (jeśli zamówione produkty są przechowywane tylko w jednym magazynie).

Bramka ALBO podczas dzielenia kieruje przepływ sekwencji tylko do jednej z gałęzi wychodzących, w oparciu o określone warunki. Podczas scalania czeka na zakończenie przepływu z tylko jednej gałęzi przychodzącej zanim będzie kontynuowała proces (von Rosing i in., 2015).

Bramka ORAZ jest używana do wykonywania dwóch lub więcej zadań, które nie mają żadnych zależności kolejnościowych od siebie i mogą być wykonywane jednocześnie (Dumas i in., 2018). Podczas scalania oczekuje na zakończenie przepływów ze wszystkich gałęzi przed kontynuowaniem procesu (von Rosing i in., 2015). Przykład użycia bramek ALBO i ORAZ pokazano na Rysunku 5.6.



Rysunek 5.6 Przykład użycia bramek ALBO i ORAZ

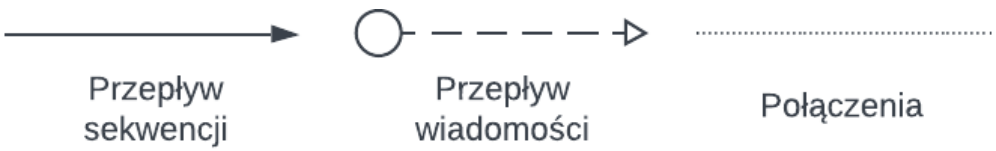
Źródło: Dumas i in. (2018)

W tym przykładzie po otrzymaniu zamówienia sprzedawca sprawdza dostępność towaru. Istnieje tylko jedna możliwa ścieżka – w zależności od tego czy produkty są w magazynie, czy nie. Z drugiej strony nie ma znaczenia, czy najpierw zostanie wykonana czynność „Wyślij fakturę", czy „Uzyskaj adres wysyłki". Ale dopiero po wykonaniu obu zestawów czynności („Uzyskaj adres wysyłki" – „Wyślij produkt" i „Wyślij fakturę" – „Odbierz płatność") zamówienie może zostać zarchiwizowane.

5.3.4. Obiekty łączące

W BPMN istnieją trzy typy obiektów łączących: przepływ sekwencji, przepływ wiadomości i połączenia.

Według von Rosing i in. (2015) **sekwencyjny przepływ** pokazuje kolejność, w jakiej zadania będą wykonywane w procesie. Jest on reprezentowany przez ciągłą linię z pełnym grotem strzałki. **Przepływ wiadomości** jest reprezentowany przez przerywaną linię. Po jednej stronie linii znajduje się okrąg, a po drugiej biały grot strzałki. Jest on używany do reprezentowania przepływu wiadomości między basenami procesów. **Połączenie** jest używane do łączenia tekstu z obiektami przepływu. Jest on reprezentowany przerywaną linię.

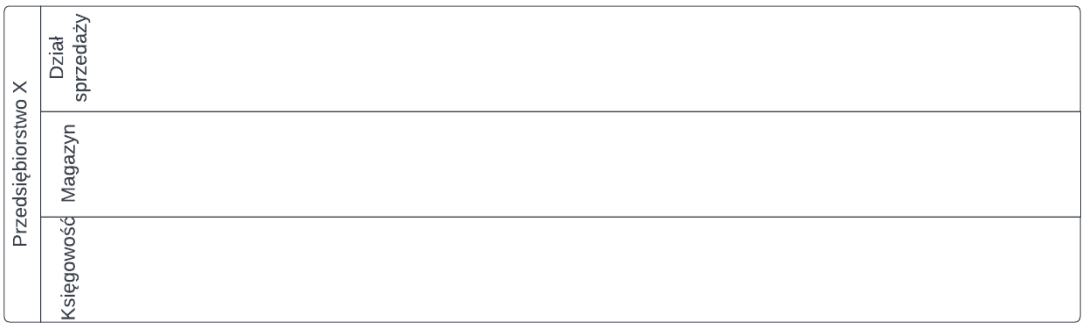


Rysunek 5.7 Przepływ sekwencji, przepływ wiadomości i połączenia

Źródło: von Rosing i in. (2015)

5.3.5. Uczestnicy

BPMN dostarcza dwa elementy do modelowania uczestników procesu: baseny i tory. Według Dumasa i in. (2018) **baseny** służą do modelowania całej organizacji, a **tory** do modelowania działu lub jednostki biznesowej. Na przykład basenem może być „Przedsiębiorstwo X", a torami „Dział sprzedaży", „Magazyn" i „Księgowość". Korzystając z basenów i torów, można łatwo zobaczyć, który uczestnik wykonuje daną czynność.



Rysunek 5.8 Baseny i tory

Źródło: Opracowanie własne

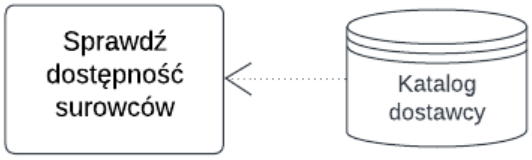
5.3.6. Artefakty

Istnieją różne typy artefaktów: obiekty danych, magazyn danych i adnotacje. **Obiekty danych** reprezentują dane, które są wymagane do wykonania pewnych zadań (dane jako dane wejściowe) lub są wynikiem wykonania zadania (dane jako dane wyjściowe). Na przykład dokument „Zamówienie" jest tworzony po wykonaniu zadania „Utwórz zamówienie". Z drugiej strony zadanie „Wyślij fakturę" wymaga faktury jako danych wejściowych, aby wykonać to zadanie. Dumas i in. (2018) twierdzą, że obiekty danych mogą być obiektami fizycznymi zawierającymi informacje (np. faktura papierowa) lub obiektami elektronicznymi (np. e-mail lub faktura w formacie PDF).

Rysunek 5.9 Obiekty danych

Źródło: Opracowanie własne

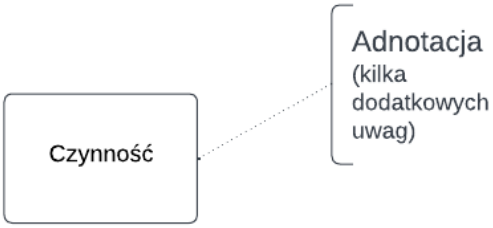
Według Dumasa i in. (2018) **magazyn danych** to miejsce, które zawiera obiekty danych, np. bazę danych dla obiektów elektronicznych lub szafkę wypełnioną obiektami fizycznymi. Magazyny danych mogą być używane przez działania procesowe do wyodrębniania i przechowywania obiektów danych. Na przykład zadanie „Sprawdź dostępność surowców" wyszukuje katalog dostawcy.



Rysunek 5.10 Magazyn danych

Źródło: Dumas i in. (2018)

Adnotacje to mechanizm, za pomocą którego twórca modelu może zapewnić odbiorcy diagramu BPMN dodatkowe informacje tekstowe (von Rosing i in., 2015).



Rysunek 5.11 Adnotacje

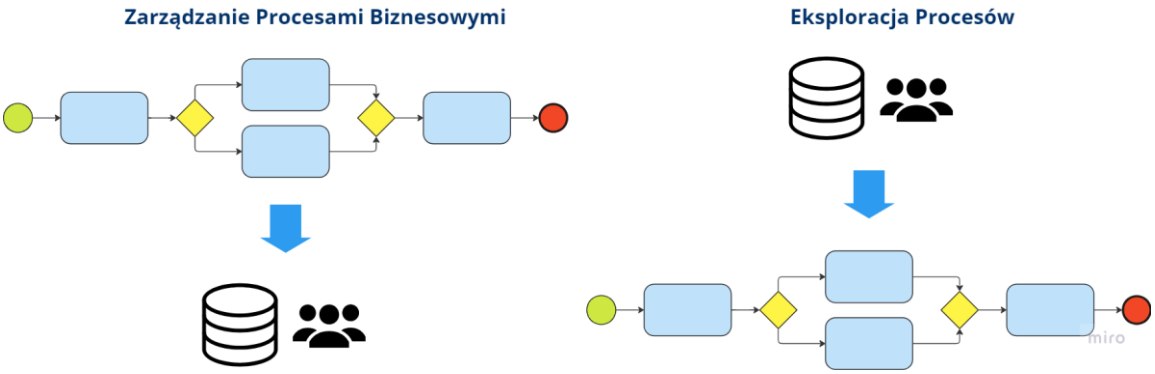
Źródło: Opracowanie własne

BPM zostało uznane za ważne ramy dla organizacji, które chcą zoptymalizować swoje operacje i dostosować swoje procesy do celów strategicznych. Ta podstawowa wiedza jest niezbędna dla zrozumienia następnego tematu.

5.4. Eksploracja procesów

Eksploracja procesów (ang. Process Mining) znajduje się na przecięciu eksploracji danych i modelowania procesów. Reprezentuje innowacyjne podejście do zrozumienia i udoskonalenia procesów biznesowych. W przeciwieństwie do teoretycznego i metodologicznego ukierunkowania BPM, eksploracja procesów bada faktyczne (rzeczywiste) dane generowane przez procesy biznesowe. Wykorzystuje dane pochodzące z różnych systemów informatycznych celem zapewnienia obiektywnego, bieżącego widoku wykonania procesu.

Rysunek 5.12. pokazuje różnicę między BPM a Process Mining. W Zarządzaniu procesami biznesowymi najpierw opracowuje się model procesu. Następnie ludzie i systemy IT wykonują zadania i czynności zgodnie z tym modelem. W Process Mining dane historyczne pochodzące z systemów IT są wykorzystywane do tworzenia modelu procesu. Model ten pokazuje faktyczne, rzeczywiste procesy.



Rysunek 5.12 Zarządzanie Procesami Biznesowymi a Eksploracja Procesów

Źródło: Opracowanie własne

IEEE (2012) definiuje **Eksplorację Procesów** jako „techniki, narzędzia i metody odkrywania, monitorowania i usprawniania rzeczywistych procesów (tj. procesów nie z góry zakładanych) poprzez wydobywanie wiedzy z dzienników zdarzeń powszechnie dostępnych w dzisiejszych systemach (informacyjnych)". Dziennik zdarzeń to cyfrowy zapis zdarzeń, które zostały wykonane w systemie informacyjnym.

Aby wykonać analizę eksploracji procesów, dziennik zdarzeń musi zawierać identyfikator przypadku, nazwę czynności i znacznik czasu. **Przypadek** (instancja procesu) to jednostka obsługiwana przez analizowany proces (np. zamówienia klientów, roszczenia ubezpieczeniowe itp.), **czynność** to dobrze zdefiniowany krok w procesie (IEEE, 2012), a **znacznik czasu** to data i godzina wykonania czynności.

Tabela 5.2. pokazuje przykład dziennika zdarzeń. W tym przykładzie występują dwa przypadki (1001 i 1002), każdy składający się z serii zdarzeń do obsługi zapytań klientów.

Tabela 5.2 Przykład dziennika zdarzeń

ID Przypadku	Nazwa czynności	Znacznik czasu	Źródło
1001	Połączenie odebrane	2023-15-12 09:00	Agent A
1001	Problem zidentyfikowany	2023-15-12 09:15	Agent A
1002	Połączenie odebrane	2023-15-12 10:17	Agent C

ID Przypadku	Nazwa czynności	Znacznik czasu	Źródło
1001	Eskalacja	2023-15-12 10:20	Agent A
1002	Podane informacje	2023-15-12 10:26	Agent C
1002	Połączenie zakończone	2023-15-12 10:28	Agent C
1001	Telefon do obsługi techn.	2023-15-12 11:43	Agent B
1001	Problem rozwiązany	2023-15-12 11:59	Agent B

Źródło: Opracowanie własne

Po wyodrębnieniu danych (dzienników zdarzeń) z systemów informatycznych (np. jako plik CSV lub XLS), dane są importowane do specjalnego oprogramowania do eksploracji procesów. Obecnie istnieje szeroki zakres oprogramowania do eksploracji procesów. Najpopularniejsze to ProM, Fluxicon Disco, ARIS Process Mining, Celonis itp. Na podstawie zaimportowanych danych oprogramowanie do eksploracji procesów odkrywa model procesu. Model ten można następnie przeanalizować w celu ustalenia, czy istnieją jakieś wąskie gardła, problemy lub możliwości udoskonalenia.

Według van der Aalst (2018) Process Mining można stosować we wszystkich rodzajach procesów operacyjnych (organizacje i systemy). Analiza procedur leczenia szpitalnego, ulepszanie procedur obsługi klienta w międzynarodowej organizacji, zrozumienie nawyków przeglądania stron przez użytkowników witryn rezerwacyjnych, ocena awarii systemu obsługi bagażu i udoskonalanie interfejsów użytkownika urządzeń rentgenowskich to tylko niektóre przykłady zastosowań.

Reil i in. (2021) przeanalizowali udaną implementację eksploracji procesów w praktycznych dziedzinach zarządzania łańcuchem dostaw. Stwierdzili, że w 2020 r. szwedzko-szwajcarska grupa technologii energetycznych i automatyzacyjnych ABB stanęła przed wyzwaniami, takimi jak połączenie ponad 40 systemów ERP i zarządzanie terabajtami danych procesowych. Implementacja eksploracji procesów w procesach produkcyjnych pozwoliła ABB uzyskać wgląd w wydajność globalnej sieci biznesowej i przejść na w pełni zdigitalizowany łańcuch dostaw. Korzyści obejmowały obniżone koszty zapasów, usprawnione procesy sprzedaży, poprawioną produktywność, terminowe dostawy, zoptymalizowane wykorzystanie sprzętu i zwiększoną wydajność. Procedury logistyki zaopatrzenia łańcucha dostaw branży automotive, które są podatne na występowanie wąskich gardeł mogących powodować duże straty przychodów, w dużym stopniu skorzystały na tej strategii. Eksploracja procesów okazała się przydatna w skutecznym rozwiązywaniu tych problemów.

Ustrukturyzowany sposób zarządzania procesami i ich usprawniania w BPM umożliwia przedsiębiorstwom dostosowanie się do zmieniających się potrzeb klientów i problemów operacyjnych. Process Mining z drugiej strony oferuje głęboki wgląd w rzeczywistą wydajność procesu, podkreślając obszary udoskonalenia. Integracja BPM i Process Mining to nie tylko strategiczna przewaga, ale także wiedza, jak ich używać i jak działają, istotne dla organizacji, aby mogły się przygotować na przyszłość.

BIBLIOGRAFIA

1. Aagesen, G. & Krogstie, J. (2015). BPMN 2.0. for Modeling Business Processes. In vom Brocke, J., Rosemann, M. (Eds.). Handbook on Business Process Management 1, 2nd edition. Heidelberg: Springer, s. 219-250.
2. Camunda (b.d.). What is Business Process Management? [dostępne: https://camunda.com/glossary/business-process-management-bpm/, dostęp 28 grudnia 2023]
3. Dumas, M., La Rosa, M., Mendling, J. & Reijers, H. A. (2018). Fundamentals of Business Process Management, 2nd Edition. Springer.
4. Freund, J. & Rücker, B. (2012). Real-Life BPMN: Using BPMN 2.0 to Analyze, Improve, and Automate Processes in Your Company. Camunda.
5. Gartner (b.d.). Business Process Management (BPM) [dostępne: https://www.gartner.com/en/information-technology/glossary/business-process-management-bpm, dostęp 28 grudnia 2023]
6. IEEE (2012). Process Mining Manifesto. IEEE Task Force on Process Mining [dostępne: https://www.tf-pm.org/upload/1580737631545.pdf, dostęp December 28, 2023]
7. Lucidchart (n.d.). What is Business Process Modeling Notation [dostępne: https://www.lucidchart.com/pages/bpmn, dostęp 28 grudnia 2023]
8. OMG (2006). Business Process Modeling Notation Specification. Object Management Group.
9. Reil, T., Groher, E. & Siegfried, P. (2021). Process Mining in Supply Chain Management. Supply Chain Management Journal, 12(2), s. 1-13.
10. Swenson, K. D. & von Rosing, M. (2015). Phase 4: What Is Business Process Management?. In von Rosing, M., Scheer, A.-W., von Scheel, H. (Eds.). The complete business process handbook. Waltham: Morgan Kaufmann, s. 79-88.
11. van der Aalst, W. (2018). Foreword: Process Mining Book. Fluxicon [dostępne: https://fluxicon.com/book/read/foreword/, dostęp 28 grudnia 2023]
12. von Rosing, M., Scheer, A.-W. & von Scheel, H. (2015). The BPM Way of Modeling. In von Rosing, M., Scheer, A.-W. and von Scheel, H. (Eds.). The complete business process handbook. Waltham: Morgan Kaufmann, s. 431-457.

13. Von Scheel, H., von Rosing, M., Fonesca, M., Hove, M. & Foldager, U. (2015). Phase 1: Process Concept Evolution. In von Rosing, M., Scheer, A.-W. and von Scheel, H. (Eds.). The complete business process handbook. Waltham: Morgan Kaufmann, s. 1-9.

6. SYSTEMY INFORMATYCZNE W LOGISTYCE

Autor: Dario Šebalj

Integracja technologii i systemów zarządzania odgrywa ważną rolę w poprawie wydajności, dokładności i podejmowania strategicznych decyzji. Trzy podstawowe systemy usprawniające operacje biznesowe w logistyce to systemy Planowania Zasobów Przedsiębiorstwa (ERP – ang. Enterprise Resource Planning), Zarządzania Magazynem (WMS – ang. Warehouse Management Systems) i Zarządzania Transportem (TMS – ang. Transportation Management Systems).

Systemy ERP stanowią kręgosłup firmy, integrując różne działy (takie jak księgowość, zaopatrzenie, sprzedaż, produkcja itp.) i procesy w ujednolicony system. Z drugiej strony WMS koncentruje się na optymalizacji operacji magazynowych, zapewniając efektywne zarządzanie zapasami i optymalizację procesów składowania. Na koniec TMS jest dedykowany do planowania, realizacji i optymalizacji transportu towarów. Wykorzystanie tego systemu jest krytyczne w redukcji kosztów transportu i poprawie efektywności logistyki.

W tym rozdziale nie tylko przedstawiono przegląd każdego systemu, ale także zbadano, w jaki sposób integracja może prowadzić do bardziej spójnego i inteligentnego środowiska biznesowego.

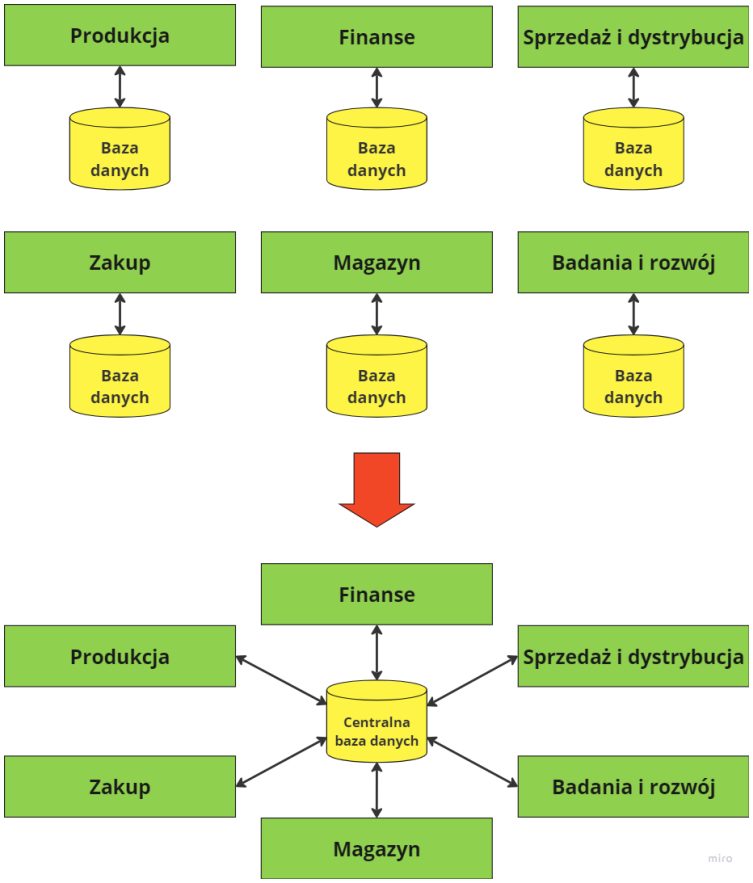
6.1. Systemy Planowania Zasobów Przedsiębiorstwa (ERP – Enterprise Resource Planning)

Na początku firmy były podzielone na różne działy, w zależności od pełnionych przez nie funkcji. Istniał więc dział produkcji, zaopatrzenia, sprzedaży, finansów itd. Każdy dział funkcjonował w izolacji, w taki sposób, że posiadał własny system gromadzenia i analizy danych. Systemy te nie były ze sobą połączone. Obecnie organizacje są uważane za jeden system, a wszystkie działy są jego podsystemami (Leon, 2014). Wszystkie współdzielą tę samą, scentralizowaną bazę danych.

Istnienie niezależnych systemów informatycznych dla każdego działu prowadziło do nieefektywności, niespójności danych i redundancji oraz wyzwań w podejmowaniu decyzji.

Przejście na zintegrowany system było więc ważnym podejściem do zarządzania organizacją, ponieważ organizacje są dzięki temu postrzegane jako pojedynczy, zunifikowany system. Krytycznym elementem tej integracji jest scentralizowana baza danych, która służy jako podstawowa część organizacji, zapewniająca wszystkim działom dostęp do spójnych **danych w czasie rzeczywistym**. Prowadzi to do lepszej komunikacji, koordynacji i współpracy między działami.

Rysunek 6.1. przedstawia różnicę między tradycyjnym podejściem, w którym działy są niezależne i każdy z nich korzysta z własnej bazy danych, a nowoczesnym podejściem, w którym działy współdzielą jedną centralną bazę danych.



Rysunek 6.1. Różnica między niezależnymi działami a działami, które korzystają z tej samej centralnej bazy danych

Źródło: Opracowanie własne na podstawie Leon (2014)

Według Bradforda (2015) systemy **planowania zasobów przedsiębiorstwa (ERP)** to systemy biznesowe, które łączą i organizują dane pochodzące z różnych działów w organizacji, celem stworzenia jednego, kompleksowego systemu, który zaspokaja potrzeby całego przedsiębiorstwa. Systemy ERP integrują, a także koordynują procesy i funkcje, które

wcześniej były rozproszone i obsługiwane przez różne starsze, niezależne systemy biznesowe, w sposób płynny, poprawiając przy tym wszystkie aspekty krytycznych operacji, w tym zakupy, księgowość, produkcję i sprzedaż.

Innymi słowy, system ERP to złożone, modułowe rozwiązanie programowe, które integruje wszystkie funkcje biznesowe firmy, pomaga w zarządzaniu procesami biznesowymi i udostępnia jedną bazę danych dla całego systemu.

System planowania zasobów przedsiębiorstwa (ERP) jest uważany za wielofunkcyjny system, który automatyzuje i integruje podstawowe operacje biznesowe organizacji w celu maksymalizacji efektywności i wydajności (Mahmood i in., 2019).

Bradford (2015) twierdzi, że firmy mogą wdrożyć moduł lub moduły oprogramowania ERP bez konieczności zakupu i wdrożenia całego pakietu, ponieważ większość z nich jest wystarczająco elastyczna.

Według Bradforda (2015) systemy ERP są często postrzegane jako systemy „back-office", ponieważ integrują funkcje zaplecza organizacyjnego, takie jak realizacja zamówień, zakupy, księgowość i finanse. Systemy ERP są obecnie czymś więcej niż tylko systemami back-office; obejmują one moduły front-office, skierowane do klienta i ułatwiające zarządzanie łańcuchem dostaw.

Systemy ERP mają wiele **zalet**, takich jak (Bradford, 2015; Paredes Hernandez, 2023):

- Poprawiona przejrzystość i przenikliwość – dane z każdego działu są dostępne dla pracowników szczebla kierowniczego;
- Dostęp do informacji w czasie rzeczywistym – dane są dostępne w czasie rzeczywistym dla wszystkich użytkowników we wszystkich działach;
- Obniżenie kosztów operacyjnych – poprzez niższe koszty zapasów, koszty produkcji lub koszty zakupów;
- Pojedynczy interfejs dla wszystkich modułów – moduły w systemie ERP wyglądają tak samo i zapewniają ten sam sposób działania;
- Skalowalność – systemy ERP oparte na chmurze umożliwiają wykorzystanie dodatkowych zasobów obliczeniowych w przypadku wzrostu firmy i danych;
- Ulepszona obsługa klienta – nowy system, taki jak oprogramowanie ERP, może umożliwić bardziej spersonalizowaną i przyspieszoną obsługę klienta, ponieważ centralizuje wszystkie dane klientów.

Niektóre z **wad** systemów ERP to (Bradford, 2015; Paredes Hernandez, 2023; Oracle, b.d.):

- Skomplikowana i czasochłonna implementacja – wdrożenie systemu ERP może trwać od kilku miesięcy do kilku lat, w zależności od wielkości firmy;
- Cena – systemy ERP są często bardzo drogie, zwłaszcza te popularne: SAP i Microsoft Dynamics NAV;
- Zarządzanie zmianą – potrzeba dużo czasu i wysiłku, aby upewnić się, że każdy ważny pracownik został odpowiednio przeszkolony w zakresie korzystania z nowego systemu.

Głównym powodem, dla którego firmy wdrażają systemy ERP, jest wspieranie rozwoju firmy. Ponadto wdrożenie ERP jest założeniem znacznej liczby firm celem zwiększenia ich produktywności i usprawnienia procesów (Software Path, 2022).

Wdrożenie systemów ERP jest bardzo skomplikowane, a większość projektów ERP kończy się niepowodzeniem. Według Saundersa (2022) około 80% projektów ERP kończy się niepowodzeniem. 25% projektów ERP zostało anulowanych lub opóźnionych, a kolejne 55% nie spełniło oczekiwań interesariuszy co do rezultatów projektu.

Mahmood i in. (2019) przeprowadzili badanie, w którym zidentyfikowali najważniejsze problemy/wyzwania, z jakimi borykają się organizacje podczas wdrażania ERP:

1. **Wsparcie najwyższego kierownictwa** – wsparcie, kierownictwo strategiczne i aktywne zaangażowanie najwyższego kierownictwa są niezbędne do pomyślnego wdrożenia i zarządzania systemami ERP.
2. **Zarządzanie zmianą** – opór, szczególnie ze strony średniego szczebla przyzwyczajonych do tradycyjnych metod, stwarza poważne wyzwania dla wdrażania nowych systemów ERP.
3. **Szkolenie i rozwój** – złożoność systemów ERP wymaga rozległego i ciągłego szkolenia pracowników, przy czym niewystarczające szkolenie prowadzi do potencjalnych awarii ERP i często stanowi ukryte koszty dla organizacji.
4. **Skuteczna komunikacja** – jasna i ciągła komunikacja oraz koordynacja między różnymi użytkownikami działów ma kluczowe znaczenie dla pomyślnego wdrożenia ERP i zarządzania zmianami organizacyjnymi.
5. **Integracja systemów** – obejmuje złożone zadanie integracji różnych modułów ERP z istniejącymi aplikacjami biznesowymi i starszymi systemami w organizacji, działanie niezbędne do optymalizacji procesów biznesowych i poprawy wydajności, ale często kosztowne i złożone.

Innym równie istotnym aspektem wdrożenia systemu ERP jest inwestycja finansowa wymagana do wdrożenia i utrzymania systemów ERP. Ten kolejny podrozdział zbada różne składniki kosztów systemu ERP, obejmujące zarówno początkową inwestycję, jak i bieżące wydatki operacyjne.

6.1.1. Koszty systemów ERP

Systemy Enterprise Resource Planning (ERP) stały się integralną częścią nowoczesnych operacji biznesowych, oferując szereg korzyści od zwiększonej wydajności po ulepszoną integrację danych. Jednakże wdrożenie takich systemów wiąże się ze znacznymi kosztami, które organizacje muszą dokładnie rozważyć.

Systemy ERP były tradycyjnie używane przez firmy sprzedające dobra materialne. Te kompleksowe oprogramowywania zostały zaprojektowane do obsługi dużych organizacji międzynarodowych. W rezultacie ich wdrożenie było niezwykle kosztowne i złożone. Moduły ERP, takie jak zakupy, sprzedaż i logistyka, stanowią podstawę procesów sprawozdawczości finansowej, a ich automatyzacja w całej organizacji globalnej może przynieść znaczne zwroty z inwestycji (Berry, 2021).

Całkowity koszt wdrożenia systemu ERP obejmuje wydatki związane z licencjonowaniem oprogramowania, wymaganiami sprzętowymi, wdrożeniem, konserwacją, konsultingiem, formalnym i nieformalnym szkoleniem i dostosowywaniem. Koszty te są zwykle określane jako **Całkowity Koszt Posiadania (TCO – ang. Total Cost of Ownership)**. Mogą się one znacznie różnić w zależności od zakresu wdrożenia, złożoności oprogramowania i wybranego dostawcy ERP. W przypadku organizacji średniej wielkości inwestycja w samo oprogramowanie ERP w pakiecie może wynieść kilka milionów dolarów (Leon, 2014; Tilley, 2020).

Oprócz kosztów oprogramowania, wdrożenie systemów ERP często wymaga znacznych inwestycji w infrastrukturę IT. Obejmuje to serwery, systemy pamięci masowej, komponenty sieciowe i ewentualnie modernizację istniejących komponentów, które zbliżają się do końca swojego cyklu życia (Bradford, 2015). Podczas gdy przetwarzanie w chmurze może obniżyć niektóre z tych kosztów sprzętowych, ponieważ oprogramowanie ERP działa na serwerach dostawcy, początkowa inwestycja w infrastrukturę pozostaje znaczącym składnikiem całkowitego kosztu.

Ukryte koszty związane z wdrożeniami ERP, takie jak opłaty za konsultacje, również odgrywają znaczącą rolę w całkowitych wydatkach. Koszty te obejmują opłaty dla zewnętrznych

konsultantów, którzy znają system ERP, ale mogą nie posiadać dogłębnej wiedzy na temat konkretnych procesów biznesowych organizacji. (Leon, 2014).

Prawie 80% całkowitych kosztów powstaje po zakupie sprzętu i oprogramowania (Tilley, 2020).

Koszt oprogramowania ERP zależy od różnych czynników. Na przykład (Hale, 2019; Wood, 2023):

- **Metoda wdrażania** – systemy ERP mogą być wdrażane w chmurze, lokalnie lub jako kombinacja obu.
- **Liczba użytkowników** – systemy ERP z mniejszą liczbą użytkowników mogą kosztować mniej.
- **Wymagane aplikacje** – liczba modułów może się wahać od modułów podstawowych do niektórych konkretnych modułów.
- **Poziom dostosowania** – wszelkie dodatkowe uaktualnienia początkowego oprogramowania zwiększają cenę systemu ERP.
- **Szkolenie i wsparcie użytkowników** – zazwyczaj, ale nie zawsze, opłaty za wdrożenie obejmują rok wsparcia technicznego dla klienta. Stałe wsparcie w czasie rzeczywistym może wiązać się z dodatkowymi kosztami.
- **Modernizacja sprzętu** – firmy muszą liczyć się z koniecznością zakupu dodatkowego sprzętu (np. serwery, pamięć masowa, infrastruktura sieciowa) podczas wdrażania, aby móc obsługiwać swój nowy system ERP.

Średni budżet na użytkownika projektu ERP, według raportu Software Path (2022), wynosi 9000 dolarów. Jednak koszt ten różni się w zależności od wielkości firmy i liczby użytkowników. Według Hale (2019) koszty utrzymania mogą wynosić od 10% do 20% początkowej opłaty licencyjnej.

6.1.2. Trendy w systemach ERP

W ostatnich dekadach organizacje wydały miliony dolarów na wdrażanie systemów ERP (Ruivo i in., 2020). Przychody z oprogramowania ERP rosną o 8% rok do roku do wartości rynkowej wynoszącej 44 miliardów dolarów w 2023 r. (Haranas, 2023) i przewiduje się, że do 2028 r. osiągną 62 miliardy dolarów (Statista, 2023).

Obecnie istnieje wielu dostawców oprogramowania ERP. Według Davidsona (2023) najlepsi dostawcy oprogramowania ERP to Microsoft, SAP, Oracle, Sage, Epicor i Infor.

Jeśli chodzi o zakup oprogramowania ERP, produkcja jest branżą o największym udziale (27%). Z wynikiem 20%, budownictwo uplasowało się na drugim miejscu. Dystrybucja i transport, które są zawarte w szerszej definicji branży łańcucha dostaw, łącznie stanowią 16% (Wood, 2023).

Według Statista (2023) **wymóg personalizacji** jest jedną z głównych preferencji klientów na rynku oprogramowania ERP. Oprogramowanie, które można dostosować do unikalnych wymagań i specyfikacji firm, jest kluczowe dla organizacji. W rezultacie wzrosło zapotrzebowanie na elastyczne i skalowalne rozwiązania ERP oparte na rozwiązaniach chmurowych. Klienci wybierają również oprogramowania, które jest proste w obsłudze i płynnie integruje się z innymi systemami.

Przyszłość i trendy systemów Planowania Zasobów Przesiębiorstwa (ERP) są kształtowane przez ewoluujące potrzeby biznesowe i postęp technologiczny. Od 2023 r. w dziedzinie ERP wyróżnia się kilka kluczowych trendów (Luther, 2023):

- **Chmurowe ERP** – Rozwiązania ERP oparte na chmurze stają się coraz bardziej popularne ze względu na prostsze wdrożenie, niższe koszty, elastyczność i możliwość dostosowania do wzrostu firmy. Pandemia przyspieszyła przejście z oprogramowania lokalnego na ERP w chmurze, ponieważ systemy te umożliwiają pracownikom łatwą pracę zdalną. Według Wood (2023) 42% firm korzystało z ERP w chmurze w 2022 r. (w porównaniu z 2013 r., kiedy odsetek ten wynosił zaledwie 4%). Zazwyczaj ERP oparte na chmurze jest dostarczane jako oprogramowanie jako usługa (SaaS – ang. Software as a Service), co oznacza, że użytkownicy muszą płacić miesięczną, kwartalną lub roczną opłatę za stały dostęp.
- **Dwupoziomowe ERP** – Podejście dwupoziomowego ERP zyskuje na popularności. Strategia ta wykorzystuje podstawowy system ERP na poziomie korporacyjnym, podczas gdy spółki zależne i działy funkcjonują przy użyciu innego, często opartego na chmurze, rozwiązania ERP. Większe firmy mogą nadal używać swojego głównego systemu ERP do celów finansowych i innych podstawowych procesów, podczas gdy mniejsze jednostki biznesowe będą szukać rozwiązań dostosowanych do ich konkretnych wymagań.
- **Transformacja cyfrowa** – systemy ERP odgrywają kluczową rolę w transformacji cyfrowej przedsiębiorstw. Integrując technologię cyfrową ze wszystkimi funkcjami

biznesowymi, systemy ERP zwiększają przychody, konkurencyjność oraz poprawiają obsługę klienta i komunikację.

- **Integracja z innymi technologiami** – nowoczesne systemy ERP są coraz częściej integrowane z innymi technologiami, takimi jak IoT i media społecznościowe, w celu usprawnienia podstawowych procesów i zapewnienia większej widoczności oraz lepszych doświadczeń klientów.
- **Personalizacja** – systemy ERP ewoluują, aby oferować klientom bardziej spersonalizowane doświadczenia, wspierane przez oparte na sztucznej inteligencji interfejsy użytkownika wspomagające i konwersacyjne, takie jak chatboty. Platformy ułatwiają wykorzystanie tego trendu poprzez wdrożenie ERP w chmurze zaprojektowane z myślą o łatwiejszej konfiguracji i rozwiązaniach branżowych.
- **Wnioski i usprawnienia oparte na sztucznej inteligencji** – sztuczna inteligencja i uczenie maszynowe są osadzane w systemach ERP, zapewniając cenne spostrzeżenia biznesowe poprzez analizę danych operacyjnych i klientów. Ta integracja pomaga w optymalizacji szeregu procesów biznesowych i poprawie personalizacji.
- **Analityka predykcyjna** – wzrost wykorzystania analityki predykcyjnej w systemach ERP poprzez koncentrację na analizie danych w celu przewidywania przyszłych trendów i wyników, co pomaga w podejmowaniu lepszych decyzji i planowaniu strategicznym.
- **Mobilne ERP** – Mobilne ERP staje się coraz bardziej powszechne, oferując dostęp do kluczowych danych biznesowych np. w podróży i ułatwiając pracę zdalną. Mobilne aplikacje ERP z przyjaznymi dla użytkownika interfejsami wspomagają wydajne realizowanie zadań przez pracowników, niezależnie od ich lokalizacji.

Trendy te wskazują na znaczącą zmianę w systemach ERP w kierunku bardziej elastycznych, spersonalizowanych i zintegrowanych rozwiązań, które są zgodne z nowoczesnymi praktykami biznesowymi i postępem technologicznym.

Systemy ERP zarządzają podstawowymi funkcjami łańcucha dostaw, takimi jak kontrola zapasów i realizacja zamówień, ale zazwyczaj są one bardzo podstawowe. Ich głównym celem było wspomaganie procesów finansowych. Moduł zarządzania zapasami ERP niezbyt dobrze sprawdził się w zarządzaniu pracą w magazynie, ale okazał się całkiem skuteczny w śledzeniu wyceny zapasów w ramach bilansu przedsiębiorstwa. W rezultacie na rynku pojawiły się najlepsze w swojej klasie aplikacje logistyczne, które mogłyby uzupełniać ERP i wypełniać powstające luki funkcjonalne. Systemy zarządzania magazynem (WMS) i systemy zarządzania

transportem (TMS) wyłoniły się jako dwie główne kategorie aplikacji logistycznych (Berry, 2021).

6.2. Systemy Zarządzania Magazynem (ang. Warehouse Management Systems)

Od momentu, w którym materiały lub towary trafiają do centrum dystrybucji lub magazynu logistyczno-przeładunkowego, aż do momentu ich opuszczenia, system zarządzania magazynem (WMS) umożliwia firmom monitorowanie i zarządzanie operacjami magazynowymi. Głównym celem WMS jest ułatwienie wydajnego i ekonomicznego przepływu materiałów i towarów przez magazyny. Pobieranie, przyjmowanie, odkładanie i śledzenie zapasów to tylko kilka z wielu zadań, które WMS wykonuje, aby ułatwić tę relokację. Systemy klasy WMS zapewniają wgląd w czasie rzeczywistym w całkowite zapasy przedsiębiorstwa, zarówno w transporcie, jak i w magazynach, a także są kluczową częścią zarządzania łańcuchem dostaw. (O'Donnell, 2020).

Według SAP (b.d.) system zarządzania magazynem optymalizuje różne działania magazynowe. Usprawnia proces przyjmowania i odkładania przy użyciu technologii RFID oraz integruje się z innym oprogramowaniem w celu wydajnej obsługi artykułów. W zarządzaniu zapasami WMS zapewnia widoczność w czasie rzeczywistym i obsługuje zaawansowaną analizę w celu lepszej kontroli zapasów. W przypadku kompletowania zamówień, pakowania i realizacji zamówień kieruje wydajnym magazynowaniem, pobieraniem i pakowaniem, wykorzystując technologie takie jak skanowanie RF (skanowanie częstotliwości radiowej – ang. radio frequency) i robotykę w celu optymalizacji przetwarzania zamówień. Procesy wysyłki są udoskonalane poprzez integrację z oprogramowaniem logistycznym, zapewniając terminowe i dokładne dostawy. WMS wspomaga również zarządzanie pracą, oferując wgląd w koszty pracy i produktywność oraz wspiera wydajne zarządzanie zadaniami. Ponadto ułatwia zarządzanie placem rozładunkowym i dokami, zwiększając wydajność załadunku, a co więcej obsługuje cross-docking dla towarów łatwo psujących się. Wreszcie WMS zapewnia cenne metryki i analizy magazynowe, umożliwiając lepsze podejmowanie decyzji i optymalizację procesów.

SAP (b.d.) wymienia 5 korzyści płynących z posiadania systemu WMS:

1. **Zwiększona wydajność operacyjna** – systemy WMS zwiększają wydajność i obsługiwane wolumeny poprzez automatyzację i usprawnienie procesów magazynowych od przyjęć po dostawy.
2. **Niższe marnotrawstwo i koszty** – WMS pomaga minimalizować marnotrawstwo, zwłaszcza w przypadku zapasów o ograniczonej dacie przydatności lub nietrwałych, a także optymalizuje wykorzystanie powierzchni magazynowej.
3. **Przejrzystość zapasów w czasie rzeczywistym** – oferuje wgląd w czasie rzeczywistym w przepływy zapasów, wspomagając dokładne prognozy popytu i poprawiając identyfikowalność.
4. **Usprawnione zarządzanie pracą** – WMS wspomaga prognozowanie potrzeb kadrowych i optymalizację alokacji zadań na podstawie różnych czynników, poprawiając tym samym morale pracowników.
5. **Lepsze relacje z klientami i dostawcami** – WMS prowadzi do lepszej realizacji zamówień i szybszych dostaw, zwiększając zadowolenie klientów i poprawiając relacje z dostawcami.

Rozwój systemów zarządzania magazynem jest nadal pod wpływem postępu technologicznego. Na przykład są to: (Scullin, 2023):

- **Zautomatyzowane narzędzia kompletacji** – technologie takie jak automatyczne głosowe kompletowanie zamówień, roboty do kompletacji zamówień i systemy pick to-light, połączone z zaawansowanymi kodami kreskowymi;
- **Roboty AGV (ang. Automated Guided Vehicles)** – usprawniają procesy składowania i pobierania, co jest kluczowe dla zadań takich jak składowanie z wykorzystaniem palet i regałów, zarządzanie kontenerami i automatyzacja procesu przyjęć;
- **Internet rzeczy (ang. Internet of Things – IoT)** – dzięki integracji w ramch IoT różnych elementów magazynu zarówno zautomatyzowanych jak i ręcznych możliwe jest kontrolowanie procesów w ramach zunifikowanej sieci, co usprawnia monitorowanie zapasów, planowanie pracy i obsługę klienta osiągnięte poprzez zwiększenie tempa realizacji zadań;
- **Rozszerzona (ang. Augmented Reality – AR) i Wirtualna Rzeczywistość (ang. Virtual Reality – VR)** – technologia AR, za pośrednictwem urządzeń takich jak inteligentne okulary, zapewnia pracownikom dostęp do specjalnych nakładek w postaci instrukcji lub informacji w czasie rzeczywistym wykorzystywanych w środowisku

magazynowym, wspomagając zadania takie jak nawigacja po trasie kompletacji czy lokalizacja pojemników bez użycia rąk. VR jest wykorzystywany do celów szkoleniowych i bezpieczeństwa, takich jak szkolenie operatorów wózków widłowych i usprawnianie tras dostaw.

Berry (2021) stwierdza, że rynek WMS jest bardzo dojrzały i istnieje wiele znanych dostawców oprogramowania, którzy oferują szeroki zakres funkcji, wspomagający nawet najbardziej skomplikowane zadania magazynowe. Wielu czołowych dostawców WMS oferuje teraz modele systemów w chmurze. 40-50% nowych klientów WMS wybiera teraz dostawę usługi w chmurze rezygnując z wdrożeń lokalnych. Do wybranych popularnych dostawców WMS zalicza się (Gartner, b.d.): SAP Extended Warehouse Management (EVM), Oracle Warehouse Management (WMS Cloud), Microsoft Dynamics 365 Supply Chain, Manhattan WMS i Infor WMS.

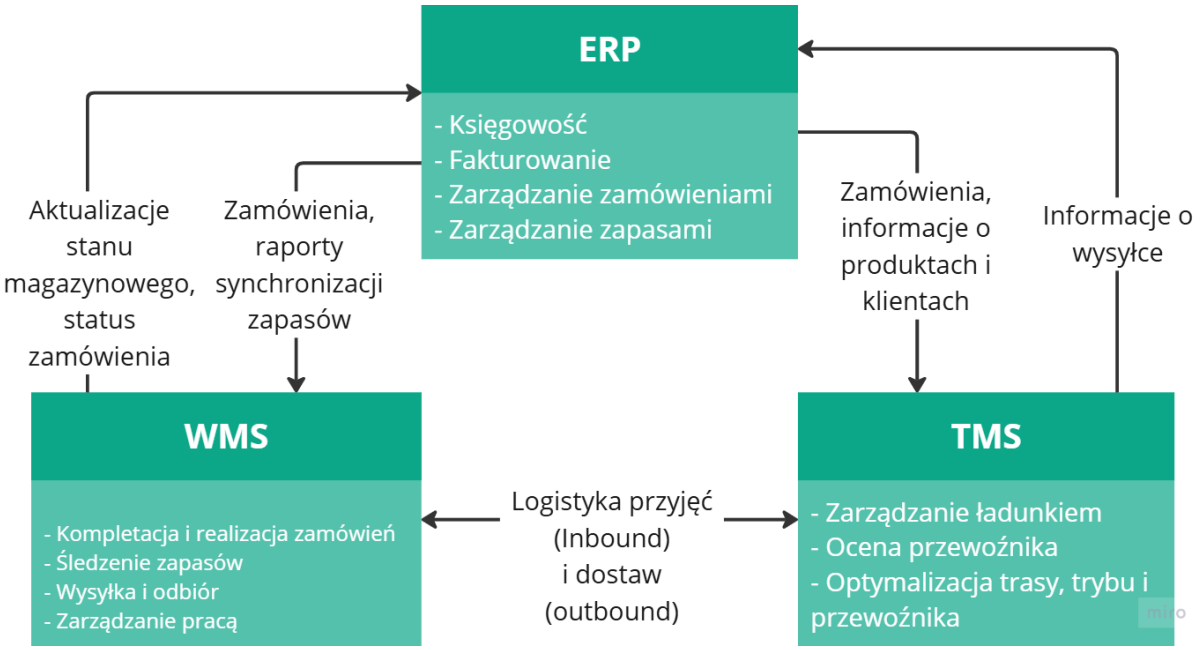
6.3. Systemy Zarządzania Transportem (ang. Transportation Management Systems)

System zarządzania transportem (TMS) to kluczowe oprogramowanie w logistyce, które optymalizuje przepływ towarów różnymi środkami transportu. Jako część szerszego systemu Zarządzania Łańcuchem Dostaw (ang. Supply Chain Management), TMS optymalizuje trasy załadunku i dostaw, śledzi ładunki i automatyzuje zadania, takie jak weryfikacja zgodności z przepisami handlowymi i rozliczanie ładunków. System ten nie tylko zapewnia terminową dostawę, ale także obniża koszty, co przynosi korzyści zarówno organizacjom, jak i ich klientom. Oferuje kompleksową widoczność operacji transportowych, wspomaga dostosowanie się do przepisów i norm, a także upraszcza proces wysyłki drogą lądową, powietrzną lub morską (SAP, b.d.; Oracle, b.d.).

Według Berry'ego (2021) istnieje kilka sposobów, w ramach których TMS może wspomóc obniżenie kosztów transportu. Dział wysyłek może zaoszczędzić dużo czasu i wysiłku automatyzując proces rezerwacji i śledzenia przesyłek. Możliwość wykorzystania asystenta planowania tras zapewnia wybór tańszej metody transportu przez pracowników działów ekspedycji. Wiele produktów TMS może optymalizować przesyłki drobnicowe (ang. Less Than Truck Load – LTL) do przesyłek pełnopojazdowych (ang. Full Truck Load – FTL), które są znacznie tańsze.

Niektóre z korzyści TMS to (SAP, b.d.; Inbound Logistics, 2023):

- **Oszczędności kosztów** – TMS znacznie obniża koszty administracyjne i transportu, optymalizując zarządzanie ładunkiem i frachtem;
- **Widoczność w czasie rzeczywistym** – zapewnia kluczowe informacje na temat procesu transportu, zwiększając wydajność trasy i śledzenie przesyłek;
- **Większe zadowolenie klienta** – zapewnia terminową dostawę i poprawia poziom obsługi klienta dzięki usprawnionym procesom śledzenia i fakturowania;
- **Poprawa wydajności** – TMS zwiększa ogólną wydajność operacji transportowych;
- **Usprawnione podejmowanie** decyzji – oferuje cenne dane do podejmowania świadomych decyzji, usprawniając strategiczne planowanie w zarządzaniu transportem.



Rysunek 6.2 Relacja pomiędzy ERP, WMS i TMS

Źródło: Essex (2020)

Rysunek 6.2. przedstawia połączenie między systemami ERP, WMS i TMS. Według Essex (2020) system ERP zarządza księgowością, fakturowaniem, zamówieniami i zarządzaniem zapasami. WMS pomaga w realizacji takich zadań jak wysyłka i przyjęcia w magazynie, kompletowanie i przechowywanie towarów, a także aktualizuje moduł inwentaryzacji systemu ERP danymi pochodzącymi ze skanów kodów kreskowych i tagów RFID w czasie rzeczywistym. System ERP dostarcza szczegóły zamówienia do TMS w celu przygotowania i wykonania wysyłki. TMS zwraca szczegóły wysyłki do ERP w celu dokonania działań księgowych i zarządzania zamówieniami oraz potencjalnie aktualizuje moduły

zarządzania relacjami z klientami (ang. Customer Relationship Management – CRM) w celu dostarczenia klientowi aktualizacji na temat statusu zamówienia.

W tym rozdziale opisano ważną rolę systemów ERP, WMS i TMS w logistyce. Systemy te, krytyczne w nowoczesnej logistyce, wspólnie zwiększają wydajność, zapewniają precyzyjne zarządzanie zapasami i optymalizują procesy transportowe. Integracja ERP, WMS i TMS to nie tylko postęp technologiczny, ale strategiczna konieczność, która prowadzi przedsiębiorstwa w kierunku większej wydajności, dokładności i satysfakcji klienta w dziedzinie zarówno logistyki jaki i zarządzania łańcuchem dostaw.

BIBLIOGRAFIA

1. Berry, J. (2021). Logistics in the Cloud-Powered Workplace. In Sullivan, M. & Kern, J. (Eds.). The Digital Transformation of Logistics. Piscataway: IEEE Press.
2. Bradford, M. (2015). Modern ERP: Select, Implement, and Use Today's Advanced Business Systems, 3rd Edition. Lulu.com
3. Davidson, R. (2023). Top 6 ERP Software Vendors. SoftwareConnect [dostępne: https://softwareconnect.com/erp/top-vendors/, dostęp 15 stycznia 2024]
4. Esex, D. (2020). Transportation management system (TMS). TechTarget [dostępne: https://www.techtarget.com/searcherp/definition/transportation-management-system-TMS, dostęp 15 stycznia 2024]
5. Gartner (b.d.). Warehouse Management Systems Reviews and Ratings [dostępne: https://www.gartner.com/reviews/market/warehouse-management-systems, dostęp 17 stycznia 2024]
6. Hale, Z. (2019). What Factors Determine the Cost of ERP Software?. Software Advice [dostępne: https://www.softwareadvice.com/resources/erp-software-pricing/, dostęp 15 stycznia 2024]
7. Haranas, M. (2023). Oracle, Microsoft, SAP, Workday Lead Cloud ERP Market: Gartner. CRN [dostępne: https://www.crn.com/news/cloud/oracle-microsoft-sap-workday-lead-cloud-erp-market-gartner, dostęp 17 stycznia 2024]
8. Inbound Logistics (2023). Transportation Management System: Meaning, Importance, and Benefits [dostępne: https://www.inboundlogistics.com/articles/transportation-management-system/, dostęp 17 stycznia 2024]
9. Leon, A. (2014). ERP demystified, 3rd edition. McGraw Hill Education.
10. Luther, D. (2023). 8 ERP Trends for 2023 & The Future of ERP. Oracle Netsuite [dostępne: https://www.netsuite.com/portal/resource/articles/erp/erp-trends.shtml, dostęp 15 stycznia 2024]
11. Mahmood, F., Khan, A. Z. & Bokhari, R. H. (2019). ERP issues and challenges: a research synthesis. Kybernetes, 49(3), pp. 629–659.
12. O'Donnell, J. (2020). Warehouse management system (WMS). TechTarget [dostępne: https://www.techtarget.com/searcherp/definition/warehouse-management-system-WMS, dostęp 15 stycznia 2024]

13. Oracle (b.d.a). What are the benefits of an ERP system? [dostępne: https://www.oracle.com/hk/erp/what-is-erp/erp-benefits/, dostęp 20 stycznia 2024]
14. Oracle (b.d.b). What Is a Transportation Management System? [dostępne: https://www.oracle.com/scm/logistics/transportation-management/what-is-transportation-management-system/, dostęp 15.01.2024]
15. Paredes Hernandez, J. (2023). The advantages and disadvantages of ERP systems. IBM [dostępne: https://ibm.com/blog/enterprise-resource-planning-advantages-disadvantages/, dostęp 16 stycznia 2024]
16. Ruivo, P., Johansson, B., Sarker, S. & Oliveira, T. (2020). The relationship between ERP capabilities, use, and value. Computers in Industry, 117, 103209.
17. SAP (b.d.a). What is a warehouse management system (WMS)? [dostępne: https://www.sap.com/products/scm/extended-warehouse-management/what-is-a-wms.html, dostęp 15 stycznia 2024]
18. SAP (b.d.b). What is a transportation management system (TMS)? [dostępne: https://www.sap.com/products/scm/transportation-logistics/what-is-a-tms.html, dostęp 15 stycznia 2024]
19. Saunders, P. (2022). Are ERP Projects Really The Stuff Of Nightmares?. Forbes [dostępne: https://www.forbes.com/sites/sap/2022/06/28/are-erp-projects-really-the-stuff-of-nightmares/, dostęp 10 stycznia 2024]
20. Scullin, Ch. (2023). 7 Smart Warehouse Technologies to Implement Today. Camcode [dostępne: https://www.camcode.com/blog/smart-warehouse-technologies/, dostęp 10 stycznia 2024]
21. Software Path (2022). What 1,384 ERP projects tell us about selecting ERP (2022 ERP report) [dostępne: https://softwarepath.com/guides/erp-report, dostęp 15 stycznia 2024]
22. Statista (2023). Enterprise Resource Planning Software – Worldwide [dostępne: https://www.statista.com/outlook/tmo/software/enterprise-software/enterprise-resource-planning-software/worldwide, dostęp 15 stycznia 2024]
23. Tilley, S. (2020). Systems Analysis and Design, 12th Edition. Boston: Cengage Learning.
24. Wood, L. (2023). How Much Does ERP Cost?. SoftwareConnect [dostępne: https://softwareconnect.com/erp/pricing/, dostęp 22 stycznia 2024]

7. E-LOGISTICS

Autor: Michał Adamczak

Rozdział ten poświęcony jest najważniejszym zagadnieniom związanym z e-logistics. Rozdział nie tylko definiuje to pojęcie ale przedstawia je także w szerszym kontekście możliwości jakie daje analiza danych w celu optymalizacji procesów logistycznych. Rozdział zawiera takie tematy jak:

- kontekst w jakim działa e-logistyka w tym koncepcję e-biznesu,
- podstawowe definicje e-logistics,
- rozwój e-logistics,
- współczesne technologie i narzędzia e-logistics,
- praktyczne rozwiazania e-logistics.

7.1. Wprowadzenie

Rozwój technologii cyfrowych ma już swoją długą historię. Nie można mówić zatem o tym, że rozwiązania cyfrowe czy dzielenie informacji w łańcuchach dostaw jest nowoczesnym rozwiązaniem. Wręcz przeciwnie z perspektywy czasu oraz punktu widzenia osób aktywnych zawodowo obecnie aspekt cyfrowy jest już rozwiązaniem dojrzałym, które na stałe wpisał się przebieg procesów logistycznych. Innymi słowy nie potrafimy już sobie wyobrazić a tym bardziej działać w logistyce bez równoległego przepływu informacji zapisanej cyfrowo.

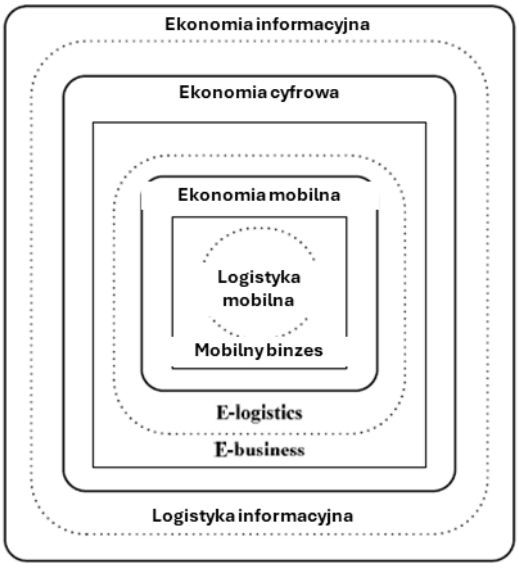
Współczesna gospodarka nazywana jest gospodarką post-industrialną lub gospodarką cyfrową. Nie oznacza to jednak, że przepływ materiałowy został całkowicie zatrzymany czy zrezygnowano z niego. To przepływ materiałowy jest kluczowy dla obrotu gospodarczego i konsumpcji. Będzie tak dopóki potrzeby ludzi będą zaspakajane przez dobra materialne. Oczywiście część potrzeb ludzi może być zaspakajana przez treści cyfrowe ale w dającej się przewidzieć przyszłości nie da się zaspokoić wszystkich potrzeb ludzi dobrami cyfrowymi. Tym samym wydaje się, że współistnienie przepływów materiałowych i cyfrowych stanowić będzie przez kolejne dekady nierozerwalny tandem. Nazywanie gospodarki cyfrową ma za zadanie jedynie podkreślić rolę i zakres w jakim współistnieją ze sobą przepływu materiałowe

i informacyjne. A co zostanie wykazane w niniejszym rozdziale przepływ cyfrowy ma coraz większe znaczenie dla budowania warunków pozwalających na poprawę efektywności przepływu materiałowego a tym samym poprawy pozycji konkurencyjnej konkretnych przedsiębiorstw jak i całych łańcuchów dostaw.

e-Logistics jest zatem rozwiązaniem wpisującym się w główny nurt współczesnej gospodarki. Jest zarazem odpowiedzią na wymagania jakie stawia współczesna gospodarka jak i rozwiązaniem dającym nowe możliwości w prowadzeniu biznesu.

7.2. E-business

E-logistics jest rozwiązaniem funkcjonującym w ramach szerszej koncepcji jaką jest e-business. E-business can be loosely defined as a business process that uses the Internet or other electronic medium as a channel to complete business transactions (Jayashankar i in., 2003). W ramach e-business można wyróżnić takie szczegółowe działania jak m.in.: e-commerce, e-advertising, e-marketing, electronic banking, electronic auctions itp. W tych rodzajach działalności gospodarczej przymiotnik „elektroniczny" wskazuje, że działania te są prowadzone wyłącznie w formie elektronicznej (cyfrowej) przy użyciu Internetu, połączenia mobilnego itp. (Skitsko, 2016). Umiejscowienie e-logistics w ramach e-business oraz innych koncepcji wykorzystujących przesyła danych przez Internet przedstawiono na rysunku 7.1.



Rysunek 7.1. E-logistics w ramach e-business

Źródło: Skitsko (2016)

W ramach e-biznesu wyróżnia się kilka podstawowych modeli komunikacji pomiędzy uczestnikami rynku (Shemet, 2012):

- B2B (business-to-business). W tym modelu występuje interakcja pomiędzy przedsiębiorstwami w celu osiągnięcia wspólnych korzyści.
- B2C (business-to-customer). W tym modelu występuje interakcja pomiędzy przedsiębiorstwami a konsumentami.
- C2C (customer-to-customer). W tym modelu następuje interakcja pomiędzy osobami fizycznymi przy wykorzystaniu różnych środków komunikacji cyfrowej.
- C2B (customer-to-business). W tym modelu opinie i pomysły klientów dotyczące produktów są przekazywane przedsiębiorstwom za pomocą różnych środków komunikacji elektronicznej np. forów, mediów społecznościowych.
- B2G (business-to-government). W tym modelu występuje interakcja pomiędzy przedsiębiorstwami a jednostkami administracji publicznej.
- C2G (customer-to-government). W tym modelu występuje interakcja pomiędzy osobami fizycznymi (obywatelami) a jednostkami administracji publicznej.
- G2B (government-to-business), G2C (government-to-customer). W tych modelach następuje interakcja pomiędzy jednostkami administracji publicznej a przedsiębiorstwami (G2B) oraz osobami fizycznymi (obywatelami (G2C).

Wpływ rozwoju technologii przetwarzania danych oraz Internetu na łańcuchy dostaw można wyróżnić w trzech obszarach (Jayashankar i in., 2003):

- rozwój systemów wspomagających zarządzanie przedsiębiorstwami (ERP) oraz planowaniem przepływów materiałowych (APS);
- rozwój systemów wspomagających proces podejmowania decyzji biznesowych, które działają w czasie rzeczywistym;
- współdzielenia informacji pomiędzy przedsiębiorstwami.

Wszystkie z wymienionych powyżej obszarów występuj również w kontekście realizowanych procesów logistycznych co stało się podstawą powstania koncepcji e-logistics.

7.3. Definicja e-logistics

Na wstępie trudno jest wykazać jedną definicję e-Logistics. Jest to spowodowane tym, że jest to pojęcie bardzo ściśle powiązane z technicznym możliwościami pozyskiwania,

gromadzenia, przetwarzania i przesyłania danych oraz informacji. Tym samym definicje tego pojęcia zmieniały się na przestrzeni czasu i zapewne dalej będą podlegały zmianom.

„E--Logistics to zbiór technologii komunikacyjnych, obliczeniowych i reguł współpracy, które przekształcają kluczowe procesy logistyczne tak, aby były zorientowane na klienta, dzieląc się danymi, wiedzą i informacjami z partnerami w łańcuchu dostaw." (Wang i in., 2004)

Inną interesującą definicję choć o zawężonym zakresie przedstawia zespół autorów Quirk, Forder and Bentley. *„E-Logistics wykorzystuje technologie internetowe do wspierania pozyskiwania: materiałów, części, wyrobów gotowych, usług magazynowania, transportu. Umożliwia także dystrybucję oraz optymalizację tras i śledzenia zapasów w łańcuchu dostaw"* (Quirk i in., 2003)

Obie przedstawiane powyżej definicje skupiają swoją uwagę na aspekcie danych jakie współtowarzyszą przepływowi materiałów w łańcuchach dostaw. Zadaniem e-logistyki jest zatem śledzenie przepływu materiałowego celem lepszej jego kontroli i dostarczania informacji o tym przepływie w czasie rzeczywistym wszystkim jego interesariuszom co w konsekwencji umożliwi synchronizację tego przepływu w łańcuchu dostaw (Mangiaracina i in., 2015).

Wg innej definicji, e-logistics to procesy logistyczne realizujące przepływ produktów kupionych w elektronicznych kanałach sprzedaży (Erceg & Damoska Sekuloska, 2019). Ilustrację takiego sposobu rozumienia e-logistics przedstawiono na rysunku.

Rysunek 7.2. E-logistics

Źródło: Moroz i in. (2014)

Zestawiając ze sobą oba podejścia do definiowania e-logistics porównano podstawowe charakterystyki logistyki tradycyjnej oraz e-logistyki wspomagającej przepływ materiałowy w handlu elektronicznym. Wyniki porównania przedstawiono w tabeli 7.1.

Tabela 7.1. Podstawowe różnice między logistyką tradycyjną a e-logistyką

Obszar	Tradycyjna logistyka	E-logistics
Typ ładunków	Palety	Paczki
Klient	Strategiczny / Znany	Nieznany
Obsługa klienta	Sztywna	Zwinna
Model dystrybucji	Push	Pull
Zapasy / Przepływ zamówień	Un-directional	Bidirectional
Miejsca docelowe	Skoncentrowane	Bardzo rozproszone
Popyt	Stabilny	Niestabilny / Seozonowy
Zamówienia	Przewidywalne	Zmienne

Źródło: Song i Hou (2004)

Podsumowując powyższe definicje należy wskazać na ich podobieństwa. Działania logistyczne realizowane na potrzeby przepływu materiałowego są do siebie bardzo zbliżone niezależnie czy dotyczą one tradycyjnego przepływu czy tego realizowanego w ramach handlu

elektronicznego. Opisując e-logistcis należy zwrócić uwagę, że zarówno w jednym jak i drugim ujęciu związane jest to pojęcie z przepływem danych opisujących przepływ materiałowy. Podstawowe funkcje e-logistics są dla obu obszarów takie same (Skitsko, 2016):

- tworzenie środowiska informacyjnego, w którym współdziałają ze sobą uczestnicy łańcucha dostaw;
- określenie charakterystyki elektronicznego przepływu informacji;
- formułowanie wymagań i potrzeb w stosunku do przedsiębiorstw, które świadczą usługi informacyjne i komunikacyjne;
- organizacja stosowania międzynarodowych standardów identyfikacji produktów;
- utrzymanie prawidłowego i niezawodnego działania, rozwój systemu informatycznego firmy;
- gromadzenie, analizowanie, przechowywanie, przekształcanie i organizacja przekazywania informacji w formie elektronicznej;
- dobór niezbędnych danych do podejmowania decyzji zarządczych.

Realizacja tych funkcji nie byłaby możliwa gdyby nie technologie cyfrowe, które umożliwiają pobieranie, gromadzenia i analizowanie danych. Opis najważniejszych z nich, które miały największy wpływ na rozwój e-logistics przedstawiono w kolejnym podrozdziale.

7.4. Rozwój e-logistics

Bazując na przedstawionych definicjach jednoznacznie można stwierdzić, ze początku e-logistics sięgają czasów kiedy powstawały pierwsze systemy informatyczne wspomagające zarządzanie przepływem materiałów, systemy material requirement planning (MRP) oraz distribution resource planning (DRP). Systemy te zaczęły się rozwijać w latach 60-tych ubiegłego wieku. Stanowiły one pierwsze rozwiązania równoległego przepływu materiałów oraz informacji zapisanej cyfrowo. Kolejne lata to dynamiczny rozwój tych systemów, który doprowadził do powstania enterprise resource planning (ERP). Równolegle rozwijały się systemy dedykowane do poszczególnych funkcji logistycznych: transport management systems (TMS) oraz warehouse management systems (WMS) (Wang, 2016). Więcej szczegółów na temat systemów informatycznych znajduje się w rozdziale 6 tego podręcznika.

Rozwój systemów klasy ERP, a szczególnie koncentracja danych oraz wielowymiarowość tych danych pozwoliła na powstanie decision support systems (DSS)

(Turbanet i in., 2002). Rozwój sieci Internet i możliwości wymiany danych pomiędzy systemami poszczególnych przedsiębiorstw zapoczątkował rozwój systemów klasy ERPII pozwalających na integrowanie danych pomiędzy partnerami w łańcuchach dostaw (Møller, 2005). Wymiana danych pomiędzy partnerami możliwa jest dzięki rozwiązaniom Electronic data interchange (EDI) (Huang i in., 2008).

Kolejnym kamieniem milowym w rozwoju e-logistcis było powstanie electronic marketplaces (EM). Powstanie platform łączących przedsiębiorstwa bezpośrednio z klientami (oraz inne konfiguracje przedstawione w podrozdziale o e-binzesie) pozwoliło na stworzenie nowych modeli biznesowych a tym samym wymagań wobec logistyki (Wang i in., 2007).

Równolegle do rozwoju EM rozwijały się systemy gromadzenia i analizy dużych zestawów danych pozwalające na realizację procesów obliczeniowych w chmurze. Rozwój technologii gromadzenia big data oraz możliwości ich analizy oraz udostępniania narzędzi analitycznych oraz wyników analiz zdalnie za pomocą sieci Interent dały całkowicie nowe możliwości szczególnie w zakresie zasilenia informacyjnego DSS a w konsekwencji możliwości optymalizacji procesów logistycznych szczególnie w takich obszarach jak: prognozowanie, zarządzanie zapasami, zarządzanie transportem oraz zarządzania zasobami ludzkimi (Waller i Fawcett, 2013).

Podsumowując rozwój technologii cyfrowych mających zastowowanie w e-logistics można się posłużyć spostrzeżeniem autorów Merali, Papadopoulos i Nadkarni (2012), którzy zaprezentowali four-step changes in ICTs w latach 60-tych, co miało duży wpływ na rozwój e-logistyki (Merali i in., 2012):

- łączność (pomiędzy ludźmi, aplikacjami i urządzeniami);
- zdolność do rozproszonego przechowywania i przetwarzania danych;
- zasięg przekazywania informacji;
- szybkość transmisji danych.

Niewątpliwie przedstawione wyżej kroki rozwoju technologii ICT wpłynęły na możliwości praktycznego zastosowania rozwiązań cyfrowych w procesach logistycznych. Zmiany te prezentują też wyraźnie kierunek w którym rozwijają się technologie cyfrowe. Technologie, które znajdują obecnie zastosowanie w e-logistics opisano szerzej w kolejnym podrozdziale.

7.5. Współczesne technologie wspierające e-logistics

Rozwój Przemysłu 4.0 i Logistyki 4.0 daje dodatkowe możliwości rozszerzenia rozwiązań i usług oferowanych w ramach e-logistics. Wśród głównych technologii wspierających e-logistics wymienia się obecnie:

- Blockchain;
- Internet Rzeczy i sensorykę (IoT);
- generatywną sztuczną inteligencję (AI);

Blockchain to rozproszony system baz danych pomiędzy wszystkimi uczestnikami tej samej sieci. System ten rejestruje i przechowuje dane w postaci połączonych ze sobą bloków tworzących zbiór rekordów. Są one trwałe i dlatego nie można ich usunąć. Ważne jest, aby wiedzieć, że nie ma możliwości dokonywania aktualizacji ani żadnych modyfikacji. Możliwe jest jednak dodanie lub odczytanie nagrania (Dutta i in., 2020).

Technologia blockchain umożliwia śledzenie różnych transakcji w całym łańcuchu dostaw w bezpieczny i identyfikowalny sposób. Udokumentowane transakcje i dane są nieodwołalnie przechowywane w łańcuchu bloków i nie mogą być używane ani odczytywane bez zgody administratora danych. Za każdym razem, gdy przesyłka jest transportowana lub przeładowywana, transakcja może zostać udokumentowana, tworząc trwałą historię od producenta do przedsiębiorcy lub konsumenta (Aritua i in., 2021).

Internet Rzeczy (IoT) pozwala na komunikowanie się ze sobą urządzeń nie będących komputerami. Koncepcja zbudowana jest w oparciu o szerokie spectrum technologii poczynając od protokołów komunikacyjnych poprzez sensory zbierające dane, infrastrukturę umożliwiającą przesyłanie danych aż po systemy analizujące zbierane dane (Minerva, 2015). Często rozwiązania IoT łączone są z sensorami RFID (radio-frequency identification) dając możliwość nie tylko lokalnej identyfikacji towarów czy ładunków ale także przekazywania tych danych do dowolnego użytkownika. Rozwiązania IoT mogą być tworzone w dwóch wariantach (Idrissi i in., 2022):

- Internet-centric – głównym elementem systemu są usługi oferowane w chmurze obliczeniowej a obiekty systemu stanowią dostawców danych;
- Object-centric – rozwiązanie, w którym centralnym punktem sieci są obiekty, którymi można sterować za pomocą komunikatów przekazywanych przez sieć Internet.

Rozwiązania IoT znajdują szerokie zastosowanie w logistyce. IoT umożliwia śledzenie różnych informacji w celu kontrolowania jakości towarów, takich jak światło, wilgotność, temperatury, wibracje, wstrząsy itp. (Dash i in., 2019) Na przykład w Maersk przewoźnik kontenerowy zamierza wprowadzić na rynek usługę, która wymaga dodatkowego ubezpieczenia na całą podróż. Warunki transportu (wibracje, temperatura, wilgotność, magnetyzm, położenie itp.) mogą być monitorowane w oprzyrządowanym kontenerze. Informacje te mogą być również przesyłane do Blockchain w celu uruchomienia częściowych płatności podczas wysyłki.

Generatywna sztuczna inteligencja to symulacja procesów inteligencji człowieka przez maszyny i systemy komputerowe. Generowanie wiedzy przez sztuczną inteligencję realizowane jest w trzech krokach (Samoili i in., 2020):

- Uczenie się - przyswajanie informacji i zasady jej wykorzystania;
- Rozumowanie - Wykorzystanie reguł do wyciągania wniosków;
- Autokorekta.

Przykładowo wykorzystanie sztucznej inteligencji pozwala na naukę opartą na historii. Pomaga to osiągnąć maksymalną wydajność, zwłaszcza w magazynach o szybkiej rotacji (takich jak magazyny e-commerce) (Dash i in., 2019).

Przedstawione technologie nie stanowią zamkniętego katalogu rozwiązań wykorzystywanych w ramach e-logistics. Szczególnie istotne jest współdziałanie tych technologii w ramach pozyskiwania, gromadzenia i przetwarzania danych w celu tworzenia informacji wspomagających podejmowanie skutecznych decyzji menedżerskich.

7.6. Praktyczne rozwiązania w zakresie e-logistics

Niezależnie o sposobu definiowania e-logistics rozwiązania te funkcjonują w zasadzie w każdym aspekcie działalności logistycznej niezależnie od realizowanej funkcji czy fazy przepływu materiałowego. Zgodnie z zaprezentowanym wcześniej przeglądem literatury uwagę należy skupić na połączeniu pomiędzy dostawcami a odbiorcami. To tu szczególnie istotna wydaje się wymiana danych i łączenie podmiotów celem poprawy efektywności

przepływu materiałowego. Jest to obecnie możliwe dzięki ogólnodostępnemu Internetowi oraz automatycznemu pozyskiwaniu danych.

Praktyczne rozwiązania w zakresie e-logistics oferują w zasadzie wszyscy operatorzy logistyczni, szczególnie Ci działający na rynku globalnym. Doskonałym przykładem rozwiązań wykorzystywanych w ramach e-logistics są te oferowane przez firmę Dachser. Ten europejski operator logistyczny udostępnia swoim klientom bezpośrednie połącznie z systemami zarządzania transportem i magazynem przez co klienci mają nieprzerwany dostęp w czasie rzeczywistym do danych na temat realizacji procesów logistycznych tego operatora. Funkcje jakie oferuje firma Dachser w ramach e-logistics to m.in. (https://www.dachser.pl/pl/elogistics-116):

- analiza produktów i usług - Dzięki temu narzędziu można szybko określić optymalne lub pożądane czasy dostawy przesyłek na terenie Europy;
- składanie zamówień online - automatyczne importowanie danych do zamówień pozwalają zaoszczędzić czas. Funkcja importu adresów z systemu ERP uzupełnia zarządzanie adresami. Funkcjonalność umożliwia także przesyłanie dokumentów, zapisywanie informacji o towarach niebezpiecznych, a także wysyłanie przyszłych zleceń oraz korzystanie z własnych kodów kreskowych;
- kontrola wszystkich kosztów transportu – umożliwia szybkie uzyskanie informacji o cenie transportu bez konieczności składania obszernych zapytań;
- śledzenie stanów magazynowych - umożliwia śledzenie procesów zachodzących w magazynach – od kontroli statusu odbioru zamówienia po monitoring partii. Funkcjonalność ta pozwala na natychmiastowe określenie braków i stanów magazynowych;
- aktualna informacja o statusie przesyłki i jej lokalizacji – funkcja Track & Trace umożliwia dla każdej przesyłki można utworzenie indywidualnego linku, który poinformuje o aktualnym statusie przesyłki. Następnie można przekazać ten link klientom lub partnerom;
- zarządzanie fakturami online - dostęp online do wszystkich danych dotyczących przesyłki. Dane dostępne są w postaci plików PDF, tabel Excel oraz plików CSV. Dane te możemy także przesłać cyfrowo poprzez centrum EDI;
- elektroniczne rozliczanie palet - zarządza sprzętem do załadunku, który wymaga śledzenia, tj. europaletami i regałami.

Kolejnym globalnym operatorem logistycznym wykorzystującym na szeroka skalę rozwiązania wywodzące się z e-logistics jest DHL. Oprócz bardzo zbliżonych funkcji zaprezentowanych powyżej dla innego operatora DHL wykorzystuje w znacznym stopniu także rozwiązania z zakresu uczenia maszynowego, rozszerzonej rzeczywistości oraz sztucznej inteligencji. Rozszerzona rzeczywistość jest wykorzystywana w celu optymalizacji infrastruktury magazynowej i prowadzonych w nim operacji logistycznych. Uczenie maszynowe i sztuczna inteligencja wykorzystywane są w celu zwiększania efektywności realizowanego biznesu oraz zwiększania odporności organizacji poprzez ukierunkowywanie podejmowanych działań na proaktywne zamiast działań reaktywnych. Podejmowanie działań proaktywnych jest możliwe dzięki analizę dużych zestawów danych i poszukiwaniu w nich związków pomiędzy przyczynami a skutkami. Możliwe jest zatem przewidywanie kształtowania się przyszłych zjawisk na podstawie przeszłych wydarzeń. Takie działania mają także wpływ na zwiększanie wartości usług kierowanych do klientów DHL i zwiększają ich pozycję konkurencyjną (DHL, 2017). Tym samym widać, że operator logistyczny może oferować nie tylko klasyczne usługi logistyczne w postaci transportu, magazynowania czy obsługi zamówień ale także zaawansowane usługi z zakresu analizy danych i rekomendowania rozwiązań płynących z tych analiz. Rozwiązania e-logistyczne stają się zatem źródłem przewagi konkurencyjnej a usługi z nich wynikające naturalnym elementem współpracy pomiędzy ogniwami łańcucha dostaw.

7.7. Podsumowanie

Rozwiązania funkcjonujące w ramach e-logistics są tak różnorodne jak definicje tego pojęcia. Można bowiem wyróżnić dwa główne nurty definiowania tego pojęcia. W szerszym ujęciu e-logistics to wszelkiego rodzaju rozwiązania cyfrowe, które towarzyszą przepływowi materiałowemu. W węższym zakresie e-logstics definiowane jest jako realizacja procesów logistycznych towarzysząca handlowi elektronicznemu. Oczywiście oba podejścia nie wykluczają się. Przedstawiona historia rozwoju, zaprezentowane rozszerzanie zakresu w jakim funkcjonuje e-logistics oraz przewidywane kierunki rozwoju jednoznacznie pokazują, że niezależnie od sposobu defniowania tego pojęcia stanowić będzie ono obiekt zainteresowania zarówno praktyków biznesu jak i badaczy.

Choć jak zaznaczono we wstępie do tego rozdziału przepływ materiałowy nie zostanie zastąpiony przez przepływ informacyjny to przepływ informacyjny w dużej mierze decyduje o efektywności przepływu materiałowego. Szczególnie istotne w tym zakresie wydaje się wspomaganie procesów informacyjnych realizowanych w ramach e-logistics metodami

i narzędziami analizy danych. Współczesne rozwiązani techniczne pozwalają na gromadzenie dużych zestawów danych oraz poszukanie związków pomiędzy tymi danymi celem przygotowania informacji przydatnych w podejmowaniu decyzji menedżerskich.

Szczegółowe rozwiązania w zakresie analizy danych, przygotowania danych do podejmowania decyzji menedżerskich zostały omówione w pozostałych rozdziałach niniejszego podręcznika. Przedstawiono w nich nie tylko koncepcje analityki biznesowej ale także systemy ERP pozwalające na gromadzenie danych, narzędzia BI pozwalające na analizę i wizualizację danych ale także nowoczesne zagadnienia związane z wykorzystaniem uczenia maszynowego w analizie danych oraz bezpieczeństwa danych.

Brak jednoznacznego kontekstu definiowania pojęcia e-logistics jest spowodowany szybkim rozwojem tematyki oraz zacieraniem się granic pomiędzy poszczególnymi rozwiązaniami wspierającymi realizację przepływu informacji.

BIBLIOGRAFIA

1. Aritua, B., Wagener, C., Wagener, N. & Adamczak, M. (2021). Blockchain solutions for international logistics networks along the new silk road between Europe and Asia, Logistics, 5(3), s. 1-14.
2. Dachser (n.d.). eLogistics: Internetowy portal do zarządzania logistyką [dostępne: dachser.pl/pl/elogistics-116, dostęp 07.04.2024]
3. Dash, R., McMurtrey, M., Rebman, C. & Kar U.K. (2019). Application of Artificial Intelligence in Automation of Supply Chain Management, Journal of Strategic Innovation and Sustainability, West Palm Beach, 14(3), pp. 43-53.
4. DHL 2017. The 21st Century Spice Trade: A Guide to the Cross-Border E-Commerce Opportunity [dostępne: http://www.dhl.com/content/dam/downloads/g0/press/publication/g0_dhl_express_cross_border_ecommerce_21st_century_spice_trade.pdf, dostęp 23.06.2018].
5. Dutta, P., Choi, T.M., Somani, S. & Butala, R. (2020). Blockchain technology in supply chain operations: applications, challenges and research opportunities. Transp Res Part E: Logist Transp Rev,142(102067).
6. Erceg, A. & Damoska Sekuloska, J. (2019). E-logistics and e-SCM: how to increase competitiveness. LogForum, 15(1), s. 155-169.
7. Huang, Z., Janz, B. & Frolick, M. (2008). A comprehensive examination of Internet-EDI adoption. Information Systems Management, 25(3), pp. 273-286.
8. Idrissi, Z. K., Lachgar, M. & Hrimech, H. (2022). Blockchain, IoT and AI revolution within transport and logistics, 2022 14th International Colloquium of Logistics and Supply Chain Management (LOGISTIQUA), EL JADIDA, Morocco, 25-27 May 2022.
9. Swaminathan, J. M. & Tayur, S. R. (2003). Models for Supply Chains in E-Business. Management Science, 49(10), s. 1387-1406.
10. Mangiaracina, R., Marchet, G., Perotti, S. & Tumino, A. (2015). A review of the environmental implications of B2C e-commerce: a logistics perspective. International Journal of Physical Distribution & Logistics Management, 45(6), s. 565-591.
11. Merali, Y., Papadopoulos, T. & Nadkarni, T. (2012). Information systems strategy: past, present, future? The Journal of Strategic Information Systems, 21(2), pp. 125–153.
12. Minerva, R., Biru A. & Rotondi, D. (2015). Towards a definition of the Internet of Things (IoT), IEEE.

13. Moroz, M., Nicu, C., Pawel, I. D. D., Polkowski, Z. (2014). The transformation of logistics into e-logistics with the example of electronic freight exchange, Zeszyty Naukowe Dolnośląskiej Wyższej Szkoły Przedsiębiorczości i Techniki. Studia z Nauk Technicznych, 3, s. 111-128.
14. Møller, C. (2005). ERP II: a conceptual framework for next-generation enterprise systems?, Journal of Enterprise Information Management, 18(4), s. 483–497.
15. Samoili, S., Cobo, M.L., Gomez, E., De Prato, G. Martinez-Plumed, F. & Delipetrev, B. (2020). Defining Artificial Intelligence. Towards an operational definition and taxonomy of artificial intelligence, Joint Research Centre, Luxembourg: Publications Office of the European Union, s. 1-97.
16. Shemet A. D. (2012). Forms of E-commerce and its place in the system of digital economy, Science and Transport Progress. Bulletin of Dnipropetrovsk National University of Railway Transport, Dnipropetrovsk, Ukraine, 41, s. 311-315.
17. Skitsko V. I. (2016). E-logistics and m-logistics in information economy. LogForum, 12(1), s. 7-16.
18. Song, Y. & Hou, H., (2004). On traditional M. F and Modern M. F, Journal of Beijing Jiaotong University (Social Sciences Edition), 3(1), s. 10-16.
19. Quirk, A., Forder, J. & Bentley, D. (2003). Electronic Commerce and the Law, 2nd edition, John Wiley & Sons Ltd., USA.
20. Turban, E., McLean, E. & Wetherbe, J. (2002). Information Technology for Management: Transforming business in the digital economy, John Wiley & Sons, New York.
21. Wang, J., Yang, D., Guo, D. & Huo Y., (2004). Taking Advantage of E-Logistics to Strengthen the Competitive Advantage of Enterprises in China [in:] Proceedings of The Fourth International Conference on Electronic Business, Bejing, s. 185-189.
22. Wang, Y., Potter, A. & Naim, M. M. (2007). Electronic marketplaces for tailored logistics, Industrial Management and Data Systems, 107 (8), s. 1170–1187.
23. Wang, Y. (2016). E-logistics: an introduction, in Wang Y.I. & Pettit S., E-Logistics: Managing Your Digital Supply Chains for Competitive Advantage,Kogan Page, s. 3-31.
24. Waller, M. A. & Fawcett, S. E. (2013). Data science, predictive analytics, and big data: a revolution that will transform supply chain design and management, Journal of Business Logistics, 34(2), s. 77–84.

8. GIS W LOGISTYCE

Autor: Dario Šebalj

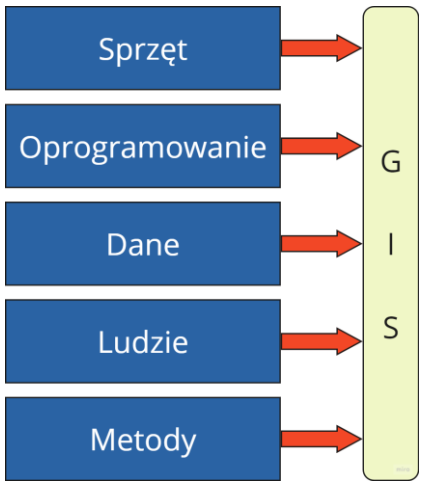
Systemy informacji geograficznej (GIS – ang. Geographic Information Systems) zrewolucjonizowały branżę logistyczną, zapewniając potężne narzędzia do analizy przestrzennej i podejmowania decyzji. Ponieważ firmy coraz częściej działają w zglobalizowanym środowisku, możliwość wizualizacji i analizy danych geograficznych jest niezbędna do optymalizacji łańcuchów dostaw, zarządzania sieciami transportowymi i zwiększania ogólnej wydajności. Technologia GIS umożliwia specjalistom ds. logistyki mapowanie tras, śledzenie przesyłek i analizowanie wzorców przestrzennych, co prowadzi do podejmowania bardziej świadomych decyzji i lepszego przydzielania zasobów. W tym rozdziale omówiono integrację GIS w logistyce, podkreślając jej zastosowania, korzyści i przyszły potencjał. Dzięki zrozumieniu, w jaki sposób GIS można wykorzystać w logistyce, firmy mogą uzyskać przewagę konkurencyjną, obniżyć koszty i zwiększyć zadowolenie klientów.

8.1. System Informacji Geograficznej (GIS)

System informacji geograficznej (GIS) to narzędzie komputerowe, które integruje, przechowuje, analizuje i wizualizuje dane geograficzne. Łączy dane przestrzenne z informacjami opisowymi, celem wspomagania użytkowników w zrozumieniu i interpretacji relacji przestrzennych, wzorców i trendów. GIS jest używany w różnych branżach do mapowania, analizy i podejmowania decyzji, zapewniając cenne informacje na temat wymiarów przestrzennych danych (Jonker, 2023; GisGeography, 2024a; Esri, b.d.; National Geographic, b.d.).

Według Esri (b.d.) i GisGeography (2024b) historia systemów informacji geograficznej (GIS) sięga początku lat 60. XX wieku, kiedy to **Roger Tomlinson**, często nazywany „ojcem GIS", opracował pierwszy skomputeryzowany GIS. Ten początkowy system został stworzony dla Canada Land Inventory, aby pomóc w zarządzaniu użytkowaniem gruntów i planowaniu zasobów. W latach 70. i 80. XX wieku postęp technologii komputerowej, teledetekcji i analizy przestrzennej doprowadził do rozwoju bardziej zaawansowanego oprogramowania GIS. W 1969 roku założono Esri (ang. Environmental Systems Research Institute), który stał się

kluczowym graczem w branży GIS, wprowadzając platformę ArcGIS. Platforma ta znacząco zwiększyła możliwości i dostępność technologii GIS. W latach 90. XX wieku technologia GIS ewoluowała, obejmując szerszy zakres zastosowań, od planowania urbanistycznego po zarządzanie środowiskiem. Integracja GIS z GPS (ang. Global Positioning Systems) i pojawienie się Internetu jeszcze bardziej rozszerzyły jej zastosowanie. Obecnie GIS jest integralnym narzędziem w różnych sektorach, w tym w transporcie, logistyce, rolnictwie i bezpieczeństwie publicznym, dostarczającym kluczowych informacji i wspomagającym podejmowanie decyzji.



Rysunek 8.1 Komponenty GIS

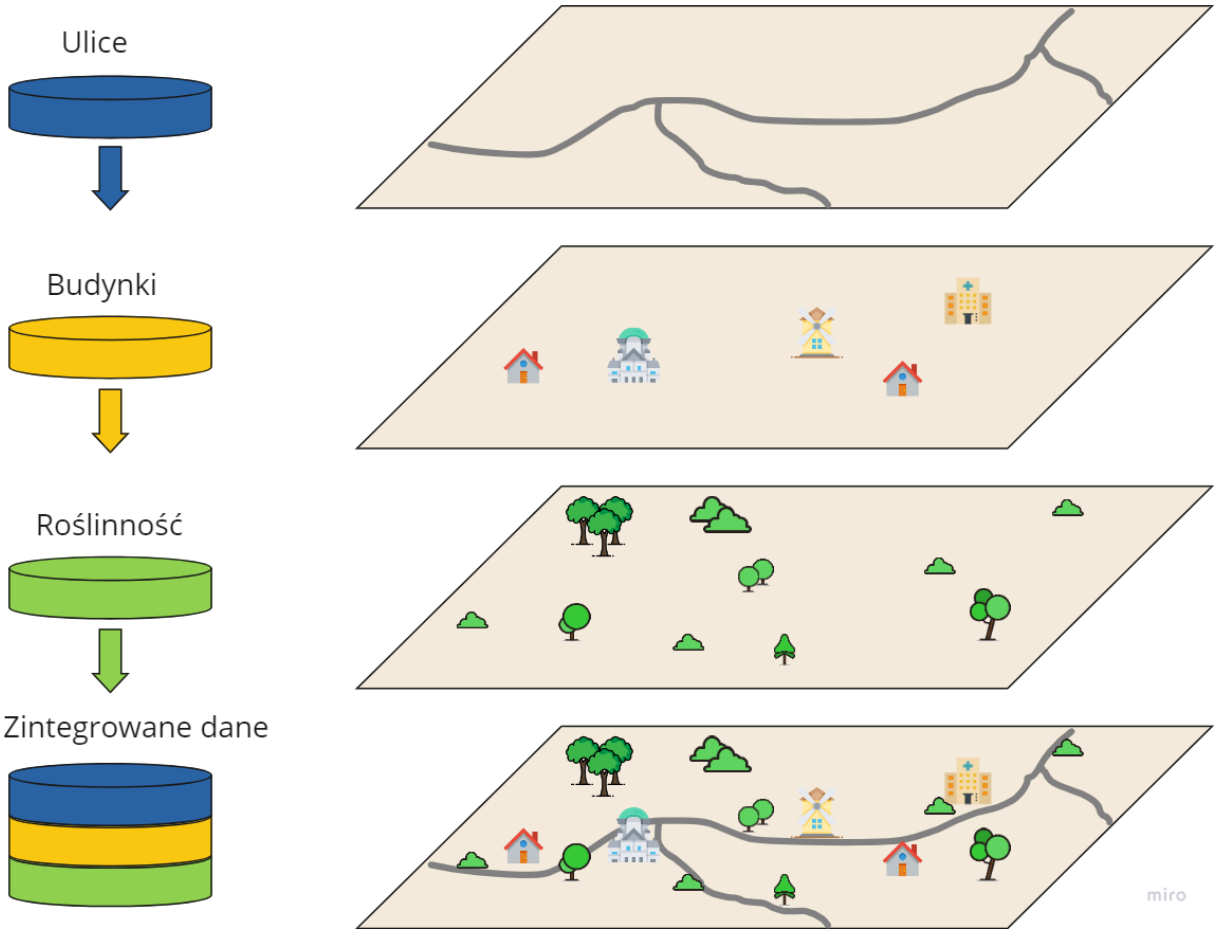
Źródło: Opracowanie własne na podstawie Kishore i Rautray (b.d.)

Rysunek 8.1 przedstawia pięć podstawowych komponentów Systemu Informacji Geograficznej (GIS) według Kishore'a i Rautray'a (b.d.):

- **Sprzęt**: Urządzenia fizyczne służące do uruchamiania oprogramowania GIS i przechowywania danych, takie jak komputery, serwery, urządzenia GPS i inne urządzenia peryferyjne.
- **Oprogramowanie**: Programy i aplikacje realizujące funkcje GIS, umożliwiające użytkownikom analizowanie i wizualizację danych przestrzennych.
- **Dane**: Informacje przestrzenne i nieprzestrzenne analizowane przez GIS, w tym mapy, zdjęcia satelitarne i dane tabelaryczne.
- **Metody**: Techniki i procedury wykorzystywane do analizy danych GIS, takie jak algorytmy i modele statystyczne.
- **Ludzie**: Profesjonaliści i użytkownicy obsługujący i zarządzający technologią GIS, od analityków danych po decydentów.

Systemy informacji geograficznej mają szeroki zakres zastosowań w różnych branżach, co czyni je niezbędnymi narzędziami do analizy danych przestrzennych i podejmowania decyzji. W analizie biznesowej GIS jest wykorzystywany do analizy rynku, wyboru lokalizacji i optymalizacji logistyki, pomagając firmom podejmować decyzje na podstawie danych w oparciu o trendy geograficzne (Longley i in., 2015). Zarządzanie środowiskiem wykorzystuje GIS do zarządzania zasobami naturalnymi, monitorowania środowiska i reagowania na katastrofy, umożliwiając skuteczniejsze działania na rzecz ochrony środowiska i planowanie awaryjne (Goodchild i in., 2018). Poprzez integrację i analizę danych przestrzennych GIS usprawnia procesy decyzyjne dzięki precyzyjnym spostrzeżeniom geograficznym i wizualizacjom, umożliwiając organizacjom identyfikację wzorców i relacji, które nie są od razu widoczne w tradycyjnych formatach danych (Longley i in., 2015).

Jednym z podstawowych komponentów technologii GIS jest koncepcja warstw. Według Esri (b.d.) warstwa to wycinek rzeczywistości geograficznej na określonym obszarze. Każda warstwa w GIS odpowiada określonemu typowi danych, takiemu jak drogi, użytkowanie gruntów, wysokość, zbiorniki wodne lub gęstość zaludnienia. Rysunek 8.2. przedstawia przykład różnych rodzajów danych na jednej mapie (ulice, budynki i roślinność), z których każdy odpowiada jednej warstwie.



Rysunek 8.2 Warstwy GIS

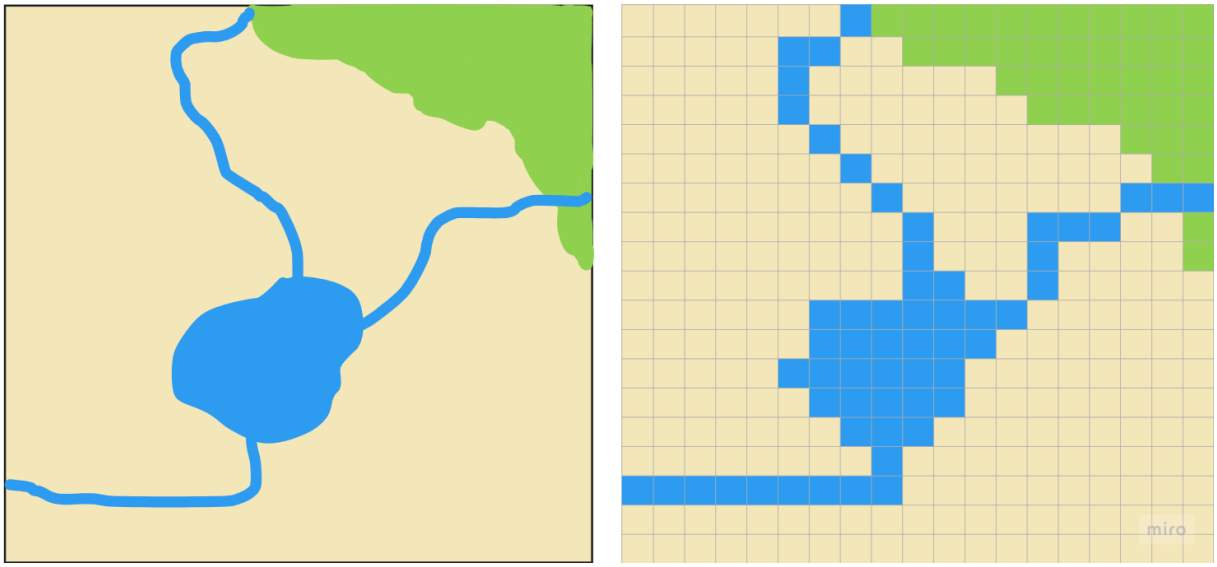
Źródło: Opracowanie własne na podstawie National Geographic (b.d.)

Systemy informacji geograficznej opierają się na różnych typach danych, aby reprezentować, analizować i wizualizować informacje geograficzne. Dane GIS można ogólnie podzielić na dwa główne typy: dane rastrowe i wektorowe.

Według Dempseya (2024) dominującą formą danych GIS są **dane wektorowe**. Punkty, linie i wielokąty używane do reprezentowania danych geograficznych są przykładami danych wektorowych. W reprezentacji wektorowej wszystkie linie są uchwycone jako punkty połączone precyzyjnie prostymi liniami (Longley i in., 2015). Dane punktowe reprezentują dyskretne punkty danych lub określone lokalizacje, takie jak szkoły, nazwy miast lub punkty zainteresowania. Dane liniowe reprezentują cechy liniowe, takie jak drogi i rzeki, a wielokąty są używane do cech powierzchniowych, takich jak jeziora, granice administracyjne i lasy. (Dempsey, 2024).

Dane rastrowe to oparta na siatce struktura danych składająca się z pikseli lub komórek, z których każda posiada powiązany atrybut. Najczęstszymi źródłami danych

rastrowych są obrazy satelitarne, zdjęcia lotnicze, dane uzyskane metodą teledetekcji oraz dane z cieniowaną rzeźbą terenu i topografią (Dempsey, 2024; Longley i in., 2015).



Rysunek 8.3 Dane wektorowe (po lewej) i rastrowe (po prawej)

Źródło: Opracowanie własne

Rysunek 8.3. przedstawia dwie reprezentacje map przy użyciu danych wektorowych (po lewej) i rastrowych (po prawej). Dane wektorowe obejmują wielokąty (jezioro i las) oraz linie (rzeki), a dane rastrowe obejmują siatkę, w której każda komórka reprezentuje jeden kolor (niebieski, żółty lub zielony).

Ważne jest zrozumienie różnych rodzajów danych GIS, aby móc je skutecznie wykorzystywać w Business Intelligence. Dane wektorowe są idealne do precyzyjnego mapowania i analizy dyskretnych cech geograficznych, podczas gdy dane rastrowe świetnie nadają się do reprezentowania danych ciągłych i danych środowiskowych na dużą skalę.

8.2. GIS w logistyce

Systemy informacji geograficznej (GIS) fundamentalnie przekształciły sektor logistyki, dostarczając narzędzi, które umożliwiają bardziej wydajne, opłacalne i strategiczne procesy podejmowania decyzji. Integracja GIS w logistyce umożliwia wizualizację, analizę i interpretację danych przestrzennych, co jest kluczowe dla optymalizacji tras, zarządzania łańcuchami dostaw i zwiększenia ogólnej wydajności operacyjnej.

Aby sprostać wyzwaniom logistycznym, technologia GIS łączy najnowocześniejsze techniki zarządzania danymi i analizy z nauką geografii. Specjaliści ds. logistyki mogą dostrzec

wzorce, relacje i trendy, które nie są widoczne w tradycyjnych formatach danych, wykorzystując geografię do łatwiejszego nakładania różnych zestawów danych na mapę. Według Esri (2017) planowanie strategiczne i optymalizacja operacyjna w dużym stopniu korzystają z tej perspektywy przestrzennej.

Jednym z głównych zastosowań GIS w logistyce jest optymalizacja tras. Analizując dane przestrzenne, firmy logistyczne mogą określić najbardziej efektywne trasy dostaw, skracając czas podróży, zużycie paliwa i ogólne koszty operacyjne. Na przykład GIS może uwzględniać wzorce ruchu, warunki drogowe, ograniczenia prędkości, aby optymalizować trasy w czasie rzeczywistym (Ramzan, 2023). Ta możliwość nie tylko zwiększa wydajność, ale także zadowolenie klientów, zapewniając terminowe dostawy.

Sureshkumar i in. (2017) przeprowadzili badanie, które uwypukla liczne zalety GIS i podkreśla jego potencjał transformacyjny w optymalizacji tras dla zarządzania ruchem drogowym. GIS umożliwia aplikację danych w czasie rzeczywistym do dynamicznych korekt ruchu drogowego i kompleksowej analizy przestrzennej, ułatwiając integrację różnych typów danych, w tym GPS i obrazów satelitarnych. Dzięki tej integracji możliwe jest zaoszczędzenie dużej ilości czasu i pieniędzy dzięki krótszym dystansom podróży i mniejszemu zużyciu paliwa. Procesy podejmowania decyzji są usprawniane dzięki możliwościom wizualizacji przestrzennej GIS, które ujawniają wzorce i trendy ukryte w konwencjonalnych formatach danych. Biorąc wszystko pod uwagę, badanie pokazuje, że optymalizacja tras oparta na GIS nie tylko zmniejsza wpływ na środowisko i zwiększa wydajność operacyjną, ale także oferuje solidne ramy do radzenia sobie ze złożonymi problemami ruchu miejskiego.

Zastosowanie systemów informacji geograficznej (GIS) w optymalizacji tras zbiórki odpadów komunalnych (MSW – ang. municipal solid waste) okazało się wysoce skuteczne w zwiększaniu wydajności operacyjnej i obniżaniu kosztów. Singh i Behera (2018) wykazali, że integracja GIS i narzędzia analityki sieci w ArcGIS przyczyniła się do znaczącego zmniejszenia odległości transportu o średnio 27,78%, co wskazuje na znaczną poprawę logistyki gospodarki odpadami w Kanpur w Indiach. Podobnie Nguyen-Trong i in. (2016) wykorzystali połączone podejście GIS, optymalizacji opartej na równaniach i modelowania agentowego celem dynamicznej optymalizacji trasy zbiórki odpadów w mieście Hagiang w Wietnamie, osiągając redukcję kosztów o 11,3%. Badania te podkreślają transformacyjny potencjał GIS w rozwiązywaniu złożoności gospodarki odpadami miejskimi, w szczególności poprzez integrację danych w czasie rzeczywistym i zaawansowanych technik modelowania. Wykorzystując GIS do analizy przestrzennej i optymalizacji tras, gminy mogą osiągnąć bardziej

zrównoważone i wydajne praktyki gospodarki odpadami, zwiększając tym samym ogólną realizację usług i zmniejszając wpływ na środowisko

Hemidat i in. (2017) przeprowadzili badanie mające na celu poprawę efektywności odbioru odpadów komunalnych stałych (MSW) w kilku miastach Jordanii przy użyciu technik GIS. Naukowcy opracowali zoptymalizowane scenariusze odbioru odpadów przy użyciu narzędzia ArcGIS Network Analyst, mające na celu zmniejszenie kosztów operacyjnych, czasu pracy pojazdów i wpływu na środowisko. Zoptymalizowane scenariusze wykazały znaczne oszczędności w porównaniu ze stanem obecnym (S0). W szczególności scenariusz S1 przyniósł oszczędności kosztów wynoszące odpowiednio 15%, 6% i 11% dla Irbidu, Karaku i Mafraku. Scenariusz S2 wykazał oszczędności kosztów wynoszące 13%, 3% i 6% dla tych samych miast. Połączony scenariusz (S3) przyniósł największe oszczędności, z redukcją całkowitych kosztów o 23%, 8% i 13%. Wyniki te podkreślają znaczący wpływ optymalizacji tras opartej na GIS na redukcję kosztów operacyjnych, czasu pracy pojazdów i wpływu na środowisko poprzez minimalizację zużycia paliwa i emisji.

Analityka oparta na GIS znacznie usprawnia zarządzanie łańcuchem dostaw krwi, zapewniając widoczność w czasie rzeczywistym i ułatwiając podejmowanie lepszych decyzji. Integracja GIS z eksploracją danych i innymi technikami analitycznymi umożliwia efektywne śledzenie, zarządzanie i optymalizację zasobów krwi, co prowadzi do poprawy wydajności operacyjnej i zmniejszenia marnotrawstwa (Delen i in., 2011).

Ponadto GIS odgrywa istotną rolę w planowaniu i zarządzaniu infrastrukturą miejską, zapewniając solidną platformę do integrowania i analizowania danych przestrzennych. Wykorzystanie GIS w tym kontekście umożliwia podejmowanie bardziej świadomych decyzji, co prowadzi do zoptymalizowanych inwestycji w infrastrukturę i ulepszonego świadczenia usług. Badanie przeprowadzone przez Irizarry i in. (2013) podkreśla skuteczność GIS w zarządzaniu infrastrukturą miejską i poprawie wydajności operacyjnej.

Wykorzystanie GIS w optymalizacji tras w różnych domenach, takich jak gospodarka odpadami komunalnymi, zarządzanie łańcuchem dostaw krwi i planowanie infrastruktury miejskiej, wykazało znaczne korzyści. GIS zwiększa wydajność operacyjną poprzez integrację danych przestrzennych z zaawansowanymi narzędziami analitycznymi, ułatwiając podejmowanie decyzji w czasie rzeczywistym i optymalizując wykorzystanie zasobów. Badania wykazały znaczną redukcję kosztów i poprawę świadczenia usług dzięki optymalizacji tras opartej na GIS, podkreślając jej krytyczną rolę w zarządzaniu złożonymi operacjami logistycznymi. Wykorzystując technologię GIS, organizacje mogą osiągnąć praktyki

zrównoważonego rozwoju, zmniejszyć wpływ na środowisko i zwiększyć ogólną skuteczność operacyjną.

8.3. Przyszłe trendy w GIS

Systemy informacji geograficznej przechodzą znaczące transformacje napędzane przez postęp technologiczny i rosnące zapotrzebowanie na analizę danych przestrzennych. Ten podrozdział zbada przyszłe trendy w GIS, skupiając się na powstających technologiach, przetwarzaniu w chmurze, integracji dużych zbiorów danych oraz roli sztucznej inteligencji (AI – ang. Artificial Intelligence) i uczenia maszynowego (ML – ang. Machine Learning).

Przyszłość GIS jest kształtowana przez kilka kluczowych trendów i innowacji, które zmieniają sposób, w jaki zbieramy, analizujemy i wykorzystujemy dane przestrzenne. Istotnym trendem jest integracja zaawansowanych technologii, takich jak przetwarzanie w chmurze, AI, uczenie maszynowe (ML) i zbieranie danych za pomocą dronów. Technologie te zwiększają wydajność i możliwości GIS, umożliwiając przetwarzanie danych w czasie rzeczywistym i bardziej zaawansowane analizy przestrzenne. Przetwarzanie w chmurze rewolucjonizuje GIS, zapewniając skalowalne i dostępne platformy do przechowywania i przetwarzania dużych zestawów danych. Ta zmiana umożliwia organizacjom wykorzystanie ogromnych ilości danych geoprzestrzennych bez potrzeby wykorzystania znacznej infrastruktury lokalnej. Wdrożenie GIS jako usługi jest coraz bardzie zauważalne, umożliwiając użytkownikom dostęp do potężnych narzędzi GIS i możliwości analizy danych za pośrednictwem platform w chmurze. Ten trend sprawia, że GIS staje się bardziej dostępny i opłacalny, szczególnie dla mniejszych organizacji i branż o ograniczonych zasobach. AI i ML odgrywają ważną rolę w automatyzacji i ulepszaniu analizy danych przestrzennych. Technologie te mogą identyfikować wzorce, tworzyć prognozy i dostarczać spostrzeżeń ze złożonych zestawów danych, których analiza ręczna byłaby trudna. Na przykład algorytmy AI mogą przetwarzać obrazy satelitarne w celu wykrywania zmian w użytkowaniu gruntów, podczas gdy modele ML mogą przewidywać wzorce ruchu na podstawie danych historycznych. Integracja AI i ML z GIS umożliwia dokładniejsze i szybsze podejmowanie decyzji w różnych sektorach, od planowania urbanistycznego po zarządzanie katastrofami. Postęp w technologii dronów jest również znaczącym trendem w GIS. Drony wyposażone w kamery i czujniki o wysokiej rozdzielczości są coraz częściej wykorzystywane do zbierania danych w trudno dostępnych obszarach. Narzędzia te zapewniają dane w czasie rzeczywistym o wysokiej dokładności, które można zintegrować z GIS w celu szczegółowego mapowania i analizy. Ten trend jest szczególnie

korzystny dla monitorowania środowiska, inspekcji infrastruktury i zarządzania rolnictwem. Innym pojawiającym się trendem jest wykorzystanie rozszerzonej rzeczywistości (AR – ang. Augmented Reality) i wirtualnej rzeczywistości (VR – ang. Virtual Reality) w GIS. Technologie te oferują nowe sposoby wizualizacji i interakcji z danymi przestrzennymi, zapewniając wciągające doświadczenia, które mogą poprawić zrozumienie i podejmowanie decyzji. Na przykład AR może nakładać dane geoprzestrzenne na widoki świata rzeczywistego, pomagając użytkownikom wizualizować podziemne instalacje lub poruszać się po złożonych środowiskach. VR może tworzyć szczegółowe symulacje krajobrazów miejskich, umożliwiając planistom eksplorację różnych scenariuszy i ich potencjalnych skutków. Analiza danych w czasie rzeczywistym staje się coraz ważniejsza w aplikacjach GIS. Możliwość przetwarzania i analizowania danych w miarę ich gromadzenia umożliwia bardziej responsywne i dynamiczne podejmowanie decyzji. Możliwości te są wzmacniane przez integrację GIS z Internetem Rzeczy (IoT – ang. Internet of Things), gdzie dane z podłączonych urządzeń mogą być stale monitorowane i analizowane. GIS w czasie rzeczywistym jest wykorzystywany w aplikacjach wykorzystywanych do takich działań jak zarządzanie ruchem drogowym, reagowanie kryzysowe i monitorowanie środowiska, gdzie terminowe informacje mają kluczowe znaczenie. Warto również zwrócić uwagę na ekspansję aplikacji GIS na nowe branże i sektory. GIS jest obecnie wykorzystywany w takich dziedzinach jak opieka zdrowotna, gdzie wspomaga śledzenie ogniska chorób i optymalizację świadczenia opieki zdrowotnej. W handlu detalicznym GIS analizuje dane demograficzne klientów i optymalizuje lokalizacje sklepów. Technologia jest również kluczowa w inicjatywach inteligentnych miast, zapewniając inteligencję przestrzenną potrzebną do efektywnego zarządzania infrastrukturą miejską i zasobami (Kerski, 2022; MGISS, 2023).

Jak widać, integracja GIS w logistyce zrewolucjonizowała branżę, zwiększając wydajność operacyjną, obniżając koszty i poprawiając zadowolenie klientów. W miarę rozwoju technologii GIS jej zastosowania w logistyce będą się rozszerzać, oferując jeszcze bardziej zaawansowane narzędzia do rozwiązywania złożonych wyzwań. Wykorzystując ten postęp, firmy logistyczne mogą utrzymać przewagę konkurencyjną i dostosować się do dynamicznych wymagań globalnego rynku.

BIBLIOGRAFIA

1. Delen, D. & Erraguntla, M. (2011). Better management of blood supply-chain with GIS-based analytics. Annals of Operations Research, 185, 181-193.
2. Dempsey, C. (2024). Types of GIS Data Explored: Vector and Raster. Geography Realm [dostępne: https://www.geographyrealm.com/geodatabases-explored-vector-and-raster-data/, dostęp 9 czerwca 2024]
3. Esri (2017). The ArcGIS Book: 10 Big Ideas about Applying The Science of Where, 2nd Edition. Esri Press.
4. Esri (b.d.a). What is GIS? [dostępne: https://www.esri.com/en-us/what-is-gis/overview, dostęp 27 maja 2024]
5. Esri (b.d.b). History of GIS? [dostępne: https://www.esri.com/en-us/what-is-gis/history-of-gis, dostęp 27 maja 2024]
6. Esri (b.d.c). Layer. GIS dictionary [dostępne: https://support.esri.com/en-us/gis-dictionary/layer, dostęp 8 czerwca 2024]
7. GisGeography (2024a). The Remarkable History of GIS [dostępne: https://gisgeography.com/history-of-gis/, dostęp 27 maja 2024]
8. GisGeography (2024b). What is GIS? Geographic Information Systems [dostępne: https://gisgeography.com/what-is-gis/, dostęp 27 maja 2024]
9. Goodchild, M. F., Steyaert, L. T., Parks, B. O., Johnston, C., Maidment, D., Crane, M. & Glendinning, S. (2018). GIS and Environmental Modeling: Progress and Research Issues, 4th Edition. John Wiley & Sons.
10. Hemidat, S., Oelgemöller, D., Nassour, A., Nelles, M. (2017). Evaluation of Key Indicators of Waste Collection Using GIS Techniques as a Planning and Control Tool for Route Optimization. Waste and Biomass Valorization, 8, 1533-1554.
11. Hguyen-Trong, K., Nguyen-Thi-Ngoc, A., Nguyen-Ngoc, D. & Dinh-Thi-Hai, V. (2017). Optimization of municipal solid waste transportation by integrating GIS analysis, equation-based, and agent-based model. Waste Management, 59, s. 14-22.
12. Irizarry, J., Karan, E. P. & Jalaei, F. (2013). Integrating BIM and GIS to improve the visual monitoring of construction supply chain management. Automation in Construction, 31, s. 241–254.

13. Jonker, A. (2023). What is a geographic information system (GIS)? IBM [dostępne: https://www.ibm.com/topics/geographic-information-system dostęp 27 maja 2024]
14. Kerski, J. (2022). 5 Trends in GIS and How to Successfully Navigate Them. Esri [dostępne: https://community.esri.com/t5/esri-young-professionals-network-blog/5-trends-in-gis-and-how-to-successfully-navigate/ba-p/1169616 dostęp 9 czerwca 2024]
15. Kishore, P. & Rautray, S. (n.d.). The five essential components of GIS. Infosys BPM [dostępne: https://www.infosysbpm.com/blogs/geospatial-data-services/gis-five-essential-components.html, dostęp 8 czerwca 2024]
16. Longley, P. A., Goodchild, M. F., Maguire, D. J. & Rhind, D. W. (2015). Geographic Information Science and Systems, 4th edition. John Wiley & Sons.
17. MGISS (2023). The Future of Gis: Trends and Innovations in Geospatial Technology [dostępne: https://mgiss.co.uk/the-future-of-gis-trends-and-innovations-in-geospatial-technology/, dostęp 9 czerwca 2024]
18. National Geographic (n.d.). GIS (Geographic Information System) [dostępne: https://education.nationalgeographic.org/resource/geographic-information-system-gis/, dostęp 27 maja 2024]
19. Ramzan, H. (2023). Optimizing Route Planning with GIS: A Comprehensive Approach for GIS Engineers. Medium [dostępne: https://medium.com/@hadiaramzan.2199/optimizing-route-planning-with-gis-a-comprehensive-approach-for-gis-engineers-f12d94dd7a16, dostęp 8 czerwca 2024]
20. Singh, S. & Behera, S. N. (2018). Development of GIS-Based Optimization Method for Selection of Transportation Routes in Municipal Solid Waste Management. Advances in Waste Management, s. 319–331.
21. Sureshkumar, M., Supraja, S. & Bhavani Sowmya, R. (2017). GIS Based Route Optimization for Effective Traffic Management. International Journal of Engineering Research And Management, 4(3), s. 62-65.

9. METODY WIZUALIZACJI DANYCH

Autor: Dario Šebalj

W dzisiejszym świecie opartym na danych, zdolność do skutecznego przekładania złożonych zbiorów danych na przejrzyste, intuicyjne wizualizacje jest koniecznością dla organizacji, które chcą efektywnie wykorzystywać swoje dane. Wizualizacja danych wykracza poza czysto estetyczną reprezentację; jest to podstawowy element analizy biznesowej, który pomaga decydentom identyfikować trendy, wartości odstające i wzorce ukryte w surowych danych. W tym rozdziale przedstawiono różne rodzaje wizualizacji, od prostych wykresów, takich jak wykresy słupkowe i liniowe, po bardziej skomplikowane reprezentacje graficzne, takie jak mapy cieplne i wykresy punktowe. Każdy typ wizualizacji służy różnym celom i jest odpowiedni dla różnych zestawów danych, więc dla analityków danych kluczowe znaczenie ma wybór odpowiedniej wizualizacji, aby skutecznie przekazać zamierzony komunikat

Wizualizacja danych jest procesem przekształcania informacji w kontekst wizualny, taki jak mapa lub wykres, i jest wykorzystywana do ułatwienia ludzkiemu umysłowi zrozumienia danych i wyciągania z nich wniosków. Głównym celem wizualizacji danych jest ułatwienie identyfikacji wzorców, trendów i wartości odstających w dużych zbiorach danych. Typowe rodzaje wizualizacji danych obejmują wykresy, tabele, mapy i pulpity nawigacyjne (ang. dashboardy) (Brush, 2022; GeeksForGeeks, 2024).

Ze względu na rosnącą popularność projektów big data i analityki danych, wizualizacja jest obecnie ważniejsza niż kiedykolwiek. Firmy coraz częściej wykorzystują uczenie maszynowe do gromadzenia ogromnych ilości danych, których przetwarzanie, zrozumienie i wyjaśnienie może być trudne i powolne. Można ten proces przyspieszyć za pomocą wizualizacji, która również ułatwia interesariuszom i właścicielom firm zrozumienie informacji (Brush, 2022). Przed wyborem metody wizualizacji ważne jest zrozumienie kontekstu wizualizacji.

9.1. Zrozumienie kontekstu sytuacyjnego

Nussbaumer Knaflic (2015) stwierdza, że zrozumienie i kontekstualizacja jest pierwszym i najważniejszym krokiem przed zaangażowaniem się w techniki wizualizacji danych

i metody opowiadania historii. Zrozumienie odbiorców jest kluczowym aspektem kontekstu. Nussbaumer Knaflic podkreśla znaczenie wiedzy o tym, kim są odbiorcy, jaki jest ich poziom wiedzy i co jest dla nich ważne. Takie zrozumienie zapewnia, że wizualizacja danych i opowiadanie historii są dostosowane do potrzeb i preferencji odbiorców, dzięki czemu informacje są bardziej istotne i angażujące.

Według IBM (b.d.), ogólne informacje podstawowe pomagają odbiorcom zrozumieć znaczenie konkretnego punktu danych. Na przykład, jeśli wskaźnik otwarć wiadomości e-mail firmy jest poniżej średniej, powinniśmy pokazać, jak wskaźnik otwarć wypada w porównaniu z całą branżą, aby zilustrować, że istnieje problem z tym kanałem marketingowym dla firmy. Odbiorcy muszą zrozumieć, jak obecna wydajność wypada w porównaniu z określonym celem, benchmarkiem lub innymi kluczowymi wskaźnikami wydajności (KPI), aby zmotywować się do podjęcia działań.

Należy więc odpowiedzieć na trzy ważne pytania (Nussbaumer Knaflic, 2015; IBM, b.d.):

- **Kto**: Obejmuje to identyfikację odbiorców i zrozumienie ich perspektywy, aby wiedzieć, jak należy dostosować historię. Takie działanie gwarantuje, że wizualizacja jest skierowana bezpośrednio do odbiorców docelowych, dzięki czemu jest bardziej skuteczna i angażująca. Na przykład, podczas gdy kwartalny raport roczny może wymagać jedynie wysokiego poziomu podsumowania danych, analityk finansowy może potrzebować szczegółowej analizy trendów na przestrzeni kilku lat. Decyzja o złożoności, poziomie szczegółowości i wglądu, który należy podkreślić, zależy od tego, kto będzie patrzył na wizualizację.
- **Co**: Odnosi się to do kluczowego komunikatu lub spostrzeżenia, które ma zostać przekazane odbiorcom. Chodzi o to, aby mieć jasność co do działania lub decyzji, na którą wizualizacja danych ma wpłynąć. Kontekst określa cel wizualizacji. Czy ma ona przekonywać, informować, badać czy potwierdzać? Każdy cel może prowadzić doróżnych decyzji dotyczących rodzaju wizualizacji i wyróżnionych punktów danych. Na przykład wizualizacja perswazyjna mająca na celu uzyskanie wsparcia dla nowej inicjatywy może koncentrować się na innych danych niż wizualizacja mająca na celu po prostu poinformowanie o dotychczasowych wynikach.
- **Jak**: Chodzi o wybranie najbardziej odpowiednich i skutecznych środków komunikowania historii lub spostrzeżeń, biorąc pod uwagę medium, format i techniki wizualizacji, które będą najlepiej rezonować z docelowymi odbiorcami. Niektóre rodzaje zestawów danych wymagają również specjalnej wizualizacji. Na przykład wykresy

punktowe są właściwe do pokazania relacji między dwiema zmiennymi, a wykresy liniowe są dobrym sposobem na pokazanie danych szeregów czasowych. Wizualizacje powinny pomóc odbiorcom zrozumieć główne przesłanie. Nieprawidłowe rozmieszczenie wykresów i danych może mieć odwrotny skutek i raczej zmylić niż oświecić odbiorców.

Jeśli chodzi o wizualizację i analizę danych, rozróżnia się wizualizację eksploracyjną i wyjaśniającą. Wizualizacja eksploracyjna motywuje użytkownika do zagłębienia się w dane lub temat w celu dokonania własnych odkryć. Wizualizacja wyjaśniająca wysuwa wyniki na pierwszy plan, przekazując czytelnikowi hipotezę lub argument autora (Schwabish, 2021).

Po zbadaniu znaczenia zrozumienia kontekstu sytuacyjnego, w którym wykorzystywane są wizualizacje danych, jasne jest, że ta podstawowa wiedza określa sposób, w jaki informacje są najlepiej przekazywane i postrzegane przez odbiorców.

Kolejnym krytycznym krokiem jest skuteczne zaangażowanie odbiorców. Następny podrozdział opisuje strategie angażowania i utrzymywania zainteresowania odbiorcy. Obejmuje to wybór elementów, które zwiększają atrakcyjność wizualną i czytelność wizualizacji danych oraz zapewniają, że kluczowe spostrzeżenia nie pozostaną niezauważone. Korzystając z technik, które przyciągają wzrok odbiorcy i podkreślają ważne dane, wizualizacje mogą być czymś więcej niż tylko informacją - mogą być wciągające i przekonujące.

9.2. Metody przyciągania uwagi

Podczas projektowania wizualizacji danych bardzo ważne jest, aby uchwycić i skierować uwagę odbiorców. Interakcja między mechaniką widzenia a zasadami percepcji wzrokowej określa, jak skutecznie wizualizacja przekazuje zamierzony komunikat. Zrozumienie, w jaki sposób ludzkie oko postrzega elementy wizualne, jest pierwszym krokiem w tworzeniu atrakcyjnych wizualizacji.

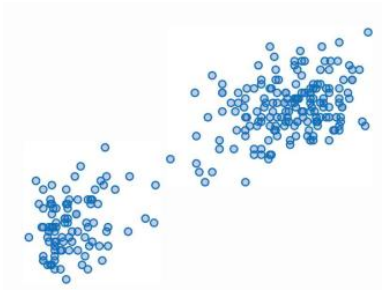
Około 70% receptorów zmysłów w naszych ciałach jest przeznaczonych do widzenia (Few, 2012).

Oko jest wstępnie przyciągane do pewnych atrybutów wizualnych, tj. atrybuty te są przetwarzane szybko i automatycznie w systemie wizualnym bez świadomego wysiłku. Atrybuty takie jak kolor, rozmiar, kształt i orientacja mogą być wykorzystane do podkreślenia

krytycznych punktów danych lub obszarów w wizualizacji i natychmiast przyciągnąć uwagę odbiorcy.

Według Schwabisha (2021) **teoria Gestalt** opisuje, w jaki sposób ludzie zazwyczaj grupują elementy wizualne. Słowo Gestalt oznacza wzór (Cairo, 2013). Została ona opracowana przez niemieckich psychologów na początku XX wieku. Jeśli chodzi o tworzenie diagramów i innych wizualizacji, sześć zasad teorii Gestalt jest szczególnie pomocnych.

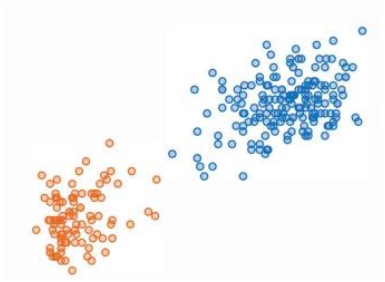
Zasada **bliskości** mówi, że nasza percepcja grupuje obiekty, gdy znajdują się one blisko siebie (np. Rysunek 9.1).



Rysunek 9.1. Bliskość jako zasada teorii Gestalt

Źródło: Schwabish (2021)

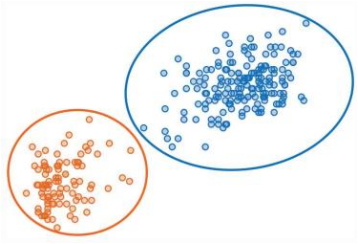
Zasada **podobieństwa** mówi, że ludzki mózg kategoryzuje obiekty na podstawie ich wspólnych atrybutów, takich jak kolor, kształt lub kierunek (np. Rysunek 9.2).



Rysunek 9.2. Podobieństwo jako zasada teorii Gestalt

Źródło: Schwabish (2021)

Zgodnie z zasadą **zamknięcia**, ograniczone obiekty są postrzegane jako grupa (np. Rysunek 9.3).



Rysunek 9.3. Zamknięcie jako zasada teorii Gestalt

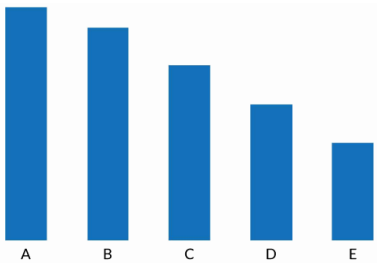
Źródło: Schwabish (2021)

Zgodnie z zasadą **domknięcia**, nasz mózg ma tendencję do ignorowania luk i uzupełniania brakujących informacji w celu utworzenia kompletnej struktury. Analizując wykres liniowy, który zawiera brakujące dane, mamy tendencję do mentalnego wypełniania luk przy użyciu najprostszego podejścia (np. Rysunek 9.4).

Rysunek 9.4. Domknięcie jako zasada teorii Gestalt

Źródło: Schwabish (2021)

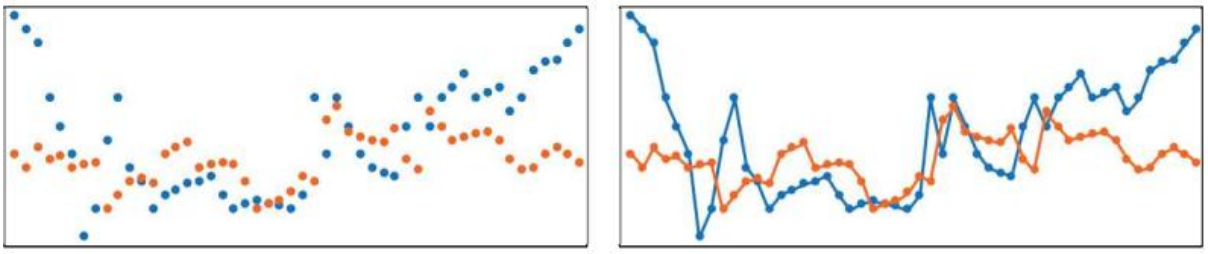
Zasada **ciągłości** sugeruje, że elementy ustawione w linii prostej lub gładkiej krzywej są postrzegane przez odbiorcę jako bardziej powiązane niż elementy, które nie leżą na linii lub krzywej (np. Rysunek 9.5).



Rysunek 9.5. Ciągłość jako zasada teorii Gestalt

Źródło: Schwabish (2021)

Opierając się na zasadzie **połączenia**, nasza percepcja kategoryzuje obiekty, które są ze sobą połączone, jako należące do tej samej grupy (np. Rysunek 9.6).



Rysunek 9.6. Połączenie jako zasada teorii Gestalt

Źródło: Schwabish (2021)

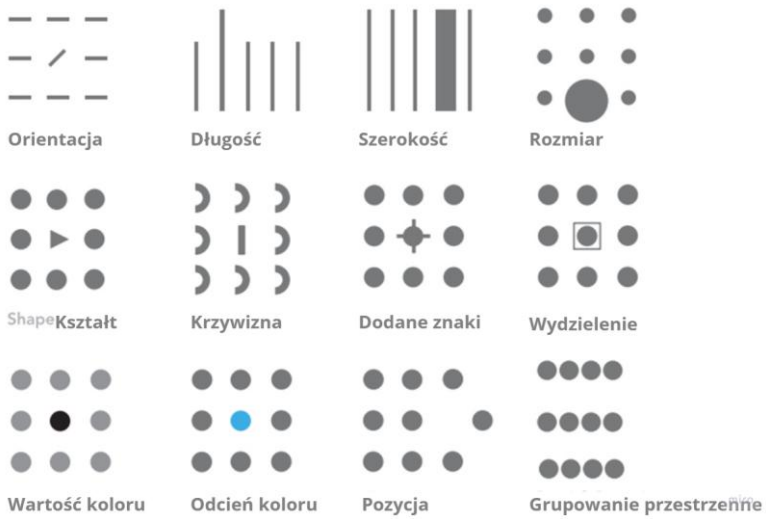
Istnieje bardzo ważna koncepcja wizualizacji danych, podzbiór teorii Gestalt zwany przetwarzaniem przeduwagowym. Schwabish (2021) wyjaśnia, że cechy przeduwagowe zwracają naszą uwagę na określony obszar wykresu lub obrazu.

Cechy te odnoszą się do jakości wizualnych, które ludzki system wizualny może postrzegać na wczesnych etapach przetwarzania wizualnego bez świadomej uwagi i które są zwykle mierzone w milisekundach. Jednym z atrybutów przeduwagowych używanych w tej książce jest kolor (niebieski) i waga (pogrubiony tekst). Kolejnym bardzo popularnym przykładem jest znalezienie określonej liczby w macierzy liczb (np. Wexler i in., 2017). Rysunek 9.7 przedstawia macierz liczb bez (po lewej stronie) i z (po prawej stronie) atrybutami przeduwagowymi.

Rysunek 9.7. Wykorzystanie atrybutów przeduwagowych w wizualizacji danych

Źródło: Schwabish (2021)

Odgadnięcie, ile jest dziewiątek na lewej macierzy zajmuje dużo czasu. Ale tylko jedna zmiana w macierzy robi dużą różnicę. Zmieniony został tylko kolor - 9 są czerwone, a wszystkie inne liczby są jasnoszare. Kolor (w tym przypadku odcień) jest jednym z kilku atrybutów uwagi wstępnej. Rysunek 9.8 pokazuje przykłady niektórych atrybutów przeduwagowych często używanych w wizualizacji danych.



Rysunek 9.8. Rodzaje atrybutów przeduwagowych używanych w wizualizacji danych

Źródło: Schwabish (2021)

Atrakcyjne atrybuty umożliwiają widzom niemal natychmiastowe rozpoznawanie wzorców, wartości odstających lub ważnych punktów danych. Stosując strategie, które kierują wzrok odbiorcy na najważniejsze informacje, wizualizacje danych mogą znacznie poprawić komunikację i zrozumienie złożonych zestawów danych.

W następnym podrozdziale przeanalizujemy, w jaki sposób różne rodzaje danych i wnioski, które mają dostarczyć, wpływają na wybór metod wizualizacji. Od prostych wykresów słupkowych po bardziej złożone mapy cieplne lub wykresy punktowe, wybór odpowiedniej metody wizualizacji ma zasadnicze znaczenie dla zapewnienia, że dane nie tylko przyciągną uwagę, ale także skutecznie i dokładnie przekażą zamierzony komunikat.

9.3. Wybór właściwej metody wizualizacji

Pierwszym krokiem w wyborze odpowiedniej metody jest dokładne zrozumienie danych. Jakie są kluczowe wiadomości, które chcemy przekazać? Z jakim typem danych mamy do czynienia? Czy pracujemy z danymi szeregów czasowych, informacjami geograficznymi lub strukturami hierarchicznymi? Rodzaj wykorzystywanych danych może mieć znaczący wpływ na wybraną metodę wizualizacji. Na przykład mapy Choropleth najlepiej nadają się do wyświetlania danych geograficznych, podczas gdy wykresy liniowe mogą być bardziej odpowiednie dla danych szeregów czasowych.

Kontekst i oczekiwania odbiorców również odgrywają ważną rolę w wyborze metody wizualizacji. Publiczność techniczna może docenić szczegółowe i złożone wizualizacje, takie jak

mapy cieplne lub diagramy sieciowe. Z drugiej strony, dla szerszej publiczności prostsze wykresy, takie jak wykresy słupkowe czy wykresy liniowe, mogą okazać się bardziej przystępne i angażujące.

Interaktywność to kolejny ważny aspekt, który należy wziąć pod uwagę. Interaktywne wizualizacje, takie jak dynamiczne dashboardy, umożliwiają użytkownikom eksplorowanie różnych poziomów danych poprzez filtrowanie, powiększanie i wybieranie określonych elementów. Ta interaktywność może prowadzić do głębszych spostrzeżeń, ponieważ użytkownicy mogą dostosować wizualizację do swoich konkretnych pytań.

Metodą wizualizacji nie zawsze muszą być wykresy. Może to być również tabela lub nawet zwykły tekst. Jak stwierdza Few (2012), celem tabel i wykresów jest skuteczne przekazywanie ważnych informacji i dostarczanie czytelnikowi ważnych, znaczących i przydatnych spostrzeżeń.

W kilku następnych podrozdziałach zostaną pokrótce wyjaśnione najpopularniejsze metody wizualizacji.

9.3.1. Prosty tekst

Nussbaumer Knaflic (2015) sugeruje użycie prostego tekstu, gdy do udostępnienia jest tylko jedna lub dwie liczby (Rysunek 9.9).

Rysunek 9.9. Prosty tekst w wizualizacji

Źródło: Schwabish (2021)

Według Schwabisha (2021) ten prosty tekst jest często określany jako BAN (ang. Big Ass Numbers). Są one zwykle używane do zwracania uwagi na kluczowe wskaźniki lub wskaźniki wydajności i zapewniają widzowi natychmiastowy dostęp do ważnych informacji. Podkreślając obszary wymagające uwagi lub działania, BAN zazwyczaj pomagają w podejmowaniu decyzji, pomagając użytkownikom skupić uwagę na ważnych aspektach danych (Tay, 2024). Chociaż BAN są proste, można je wzbogacić o subtelne elementy wizualne, takie jak kodowanie kolorami lub ikony, aby wskazać wydajność w stosunku do celów

lub zmian w czasie. Na przykład czerwona strzałka w dół obok liczby sprzedaży może natychmiast wskazywać spadek, podczas gdy zielona strzałka w górę sygnalizuje wzrost.

9.3.2. Tabela

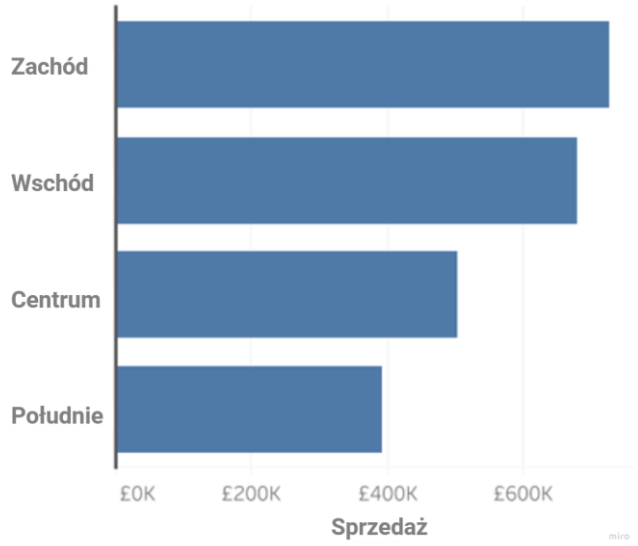
Tabele są istotną częścią wizualizacji danych, ponieważ zapewniają uporządkowany i przejrzysty sposób prezentacji danych liczbowych. Tabele są niezwykle przydatne, jeśli chodzi o prezentowanie szczegółowych informacji z precyzją i przejrzystością, mimo że mogą nie mieć takiego samego wpływu wizualnego jak wykresy lub mapy. Według Schwabisha (2020), w większości przypadków nie mają one na celu zapewnienia szybkiej wizualnej reprezentacji danych. Zamiast tego tabele są przydatne, gdy konieczne jest pokazanie dokładnych wartości danych lub szacunków. Chociaż nie są one idealną opcją do prezentacji dużej ilości danych lub na małej przestrzeni, dobrze zaprojektowana tabela może pomóc czytelnikowi znaleźć określone liczby, a także dostrzec trendy i wartości odstające. Few (2012) zauważa, że tabele są użyteczne dla odniesień i porównań jeden do jednego ze względu na ich prostą strukturę i fakt, że wartości ilościowe są wyrażone jako tekst, który możemy natychmiast zrozumieć bez konieczności tłumaczenia go.

Tabele powinny być sformatowane w następujący sposób (Schwabish, 2020; Nussbaumer Knaflic, 2015):

- usuń wszystkie obramowania wokół tabeli;
- rozjaśnij linie siatki tak bardzo, jak to możliwe lub usuń je całkowicie;
- wyraźnie oddziel nagłówek od treści tabeli;
- wyrównaj tekst w tabeli i nagłówku do lewej strony, a liczb do prawej;
- używaj odpowiedniego poziomu szczegółowości danych (np. używaj liczb z jednym miejscem po przecinku, jeśli jest to wystarczające do zrozumienia danych).

9.3.3. Wykres słupkowy

Wykres słupkowy jest idealny do wyświetlania wartości liczbowych według grup lub kategorii (np. jeśli chcemy wyświetlić liczbę pracowników według działu). Może być wyrównany pionowo lub poziomo. Wyrównanie poziome (jak na Rysunku 9.10) jest zalecane, jeśli nazwy kategorii są zbyt długie lub jeśli istnieje zbyt wiele kategorii. Wykres słupkowy na rysunku 9.10. przedstawia sprzedaż (dane ilościowe) w podziale na regiony (dane jakościowe).



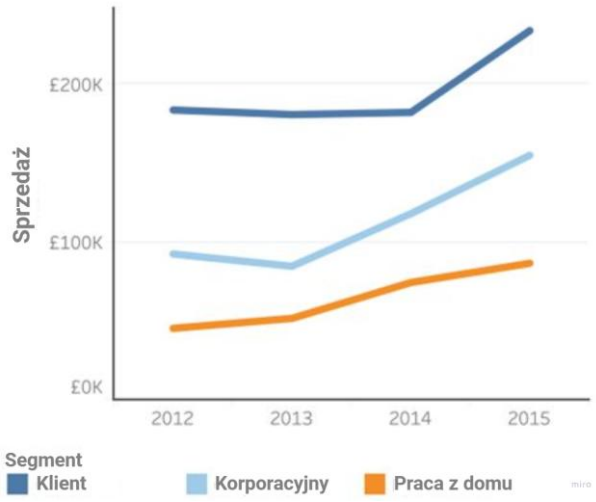
Rysunek 9.10. Wykres słupkowy w wizualizacji

Źródło: Wexler i in. (2021)

Według Few (2013) najskuteczniejszym sposobem przedstawienia miar związanych z dyskretnymi pozycjami na skali nominalnej lub porządkowej jest wykres słupkowy. Łatwo jest porównać poszczególne wartości, po prostu porównując wysokość słupków. Oś wykresu słupkowego musi zaczynać się od zera. Jeśli oś zaczyna się od wartości innej niż zero, może to nadmiernie uwypuklić różnicę między słupkami i zniekształcić nasze postrzeganie wartości na wykresie słupkowym, które opiera się na długości słupków (Schwabish, 2021).

9.3.4. Wykresy liniowe

Wykres liniowy służy do przedstawienia zmiany wartości ilościowej, która leży na osi y, w czasie, który jest umieszczony na poziomej osi x. Yi (b.d.a) sugeruje, że wykres liniowy nie powinien zawierać więcej niż pięć linii. Ponadto nie jest konieczne uwzględnianie zerowej linii bazowej dla wykresu liniowego. Dopuszczalne jest rozszerzenie zakresu osi pionowej do punktu, w którym zmiany wartości są najbardziej pouczające, jeśli linia zerowa nie jest ani zrozumiała, ani pomocna.

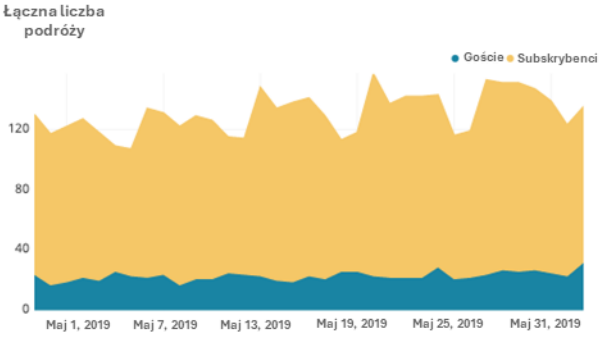


Rysunek 9.11. Wykres liniowy w wizualizacji

Źródło: Wexler i in. (2021)

Wykres liniowy na Rysunku 9.11 przedstawia sprzedaż (dane ilościowe) w okresie 4 lat i jest również podzielony na segmenty.

Wykres obszarowy (Rysunek 9.12), który jest wariantem wykresu liniowego, dodaje cień między linią a zerową linią bazową (Yi, b.d.a).



Rysunek 9.12. Wykres obszaru w wizualizacji

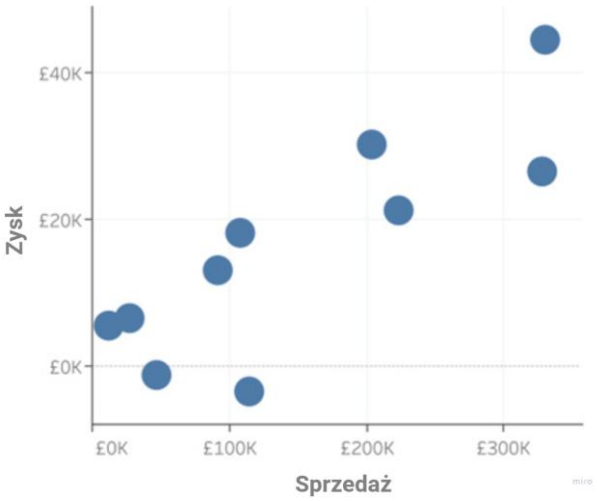
Źródło: Yi (b.d.)

Wykres obszarowy można postrzegać jako skrzyżowanie wykresu liniowego i słupkowego, ponieważ wartości można interpretować nie tylko na podstawie ich pozycji w pionie, ale także na podstawie obszaru zacienionego między każdym punktem a linią bazową (Yi, n.d.).

9.3.5. Wykres punktowy

Wykres punktowy jest używany do sprawdzenia, czy istnieje związek między dwiema zmiennymi ilościowymi. Według The Data Visualisation Catalogue (b.d.) wzorce widoczne na

wykresie punktowym można wykorzystać do interpretacji charakteru korelacji. Są to wartości: dodatnie (wartości rosną razem), ujemne (jedna wartość maleje, podczas gdy druga rośnie) lub zero (brak korelacji).



Rysunek 9.13. Wykres punktowy w wizualizacji

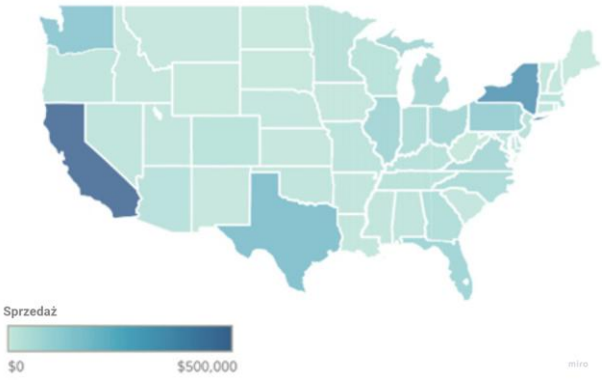
Źródło: Wexler i in. (2017)

Wykres punktowy na Rysunku 9.13 pokazuje związek między zyskiem a sprzedażą (obie zmienne ilościowe).

Według Yi (b.d.) bardzo ważne jest, aby wspomnieć, że na wykresie punktowym, nawet jeśli widzimy związek między dwiema zmiennymi, nie oznacza to, że zmiany jednej zmiennej powodują zmiany drugiej. Prowadzi to do powszechnie używanego w statystyce wyrażenia: „korelacja nie implikuje związku przyczynowego".

9.3.6. Kartogram

Kartogram wykorzystuje różnice w cieniowaniu lub kolorowaniu w obrębie wcześniej zdefiniowanych obszarów, aby wskazać wartości lub kategorie w tych obszarach (Wexler i in., 2017). Według Schwabisha (2021) paleta kolorów na kartogramie jest łatwa do zrozumienia, mniejsze wartości odpowiadają jaśniejszym kolorom, a większe wartości ciemniejszym kolorom.



Rysunek 9.14. Kartogram w wizualizacji

Źródło: Wexler i in. (2017)

Kartogram na Rysunku 9.14 pokazuje sumę sprzedaży w różnych stanach USA.

9.3.7. Mapa cieplna

Mapa cieplna to wizualizacja danych w formacie tabelarycznym, gdzie kolorowe komórki reprezentują względną wielkość liczb (Nussbaumer Knaflic, 2015).

Ponieważ kolor jest kluczowym elementem tego typu wykresu, należy upewnić się, że wybrana paleta kolorów pasuje do danych. Najpopularniejszym rodzajem koloru jest kolor sekwencyjny, w którym ciemniejsze kolory korelują z wyższymi wartościami, a jaśniejsze kolory korelują z niższymi wartościami lub odwrotnie (Yi, b.d.b).

	Region A	Region B	Region C
Kategoria 1			
Kategoria 2			
Kategoria 3			
Kategoria 4			
Kategoria 5			

Rysunek 9.15. Mapa cieplna w wizualizacji

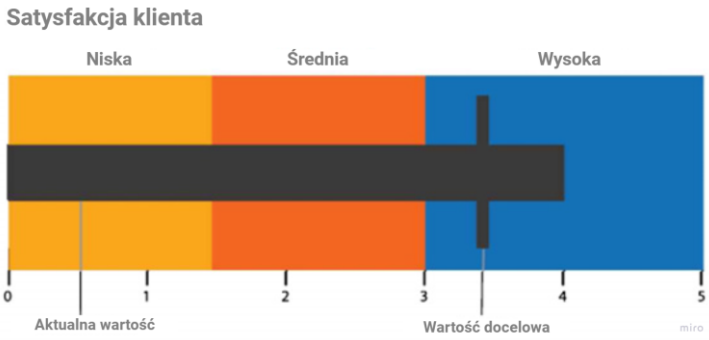
Źródło: Opracowanie własne na podstawie Nussbaumer Knaflic (2015)

Mapa cieplna na Rysunku 9.15 pokazuje różne wartości niektórych danych ilościowych (np. sprzedaży) według kategorii (w wierszach) i regionu (w kolumnach).

9.3.8. Wykres pociskowy

Wykres pociskowy został wynaleziony przez Stephena Few w 2005 roku (Few, 2013). Jest to w zasadzie wykres słupkowy z pojedynczym czarnym poziomym paskiem

reprezentującym rzeczywistą wartość, dodatkowym (pionowym) znacznikiem wartości docelowej (którą chcemy osiągnąć) i zacienionymi obszarami w tle reprezentującymi skalę sukcesu (np. słaby, dobry, doskonały).



Rysunek 9.16. Wykres pociskowy w wizualizacji

Źródło: Schwabish (2021)

Wykres pociskowy na Rysunku 9.16 pokazuje, że chcemy osiągnąć ocenę zadowolenia klienta na poziomie 3,4 (w skali 5). Nasza obecna ocena zadowolenia wynosi 4, co jest wartością wyższą od wartości docelowej. W tle znajdują się trzy obszary zadowolenia klienta – niski (słaby), średni (dobry) i wysoki (doskonały).

Wybór odpowiedniej metody wizualizacji jest bardzo ważny dla skutecznej komunikacji, ale zasady projektowania, które kierują tymi technikami, również odgrywają kluczową rolę w przejrzystości i skuteczności prezentacji danych. W następnej sekcji przyjrzymy się znaczeniu układu, typografii, schematów kolorów i strategicznego wykorzystania przestrzeni, które mają kluczowe znaczenie dla uczynienia wizualizacji nie tylko estetycznymi, ale także łatwymi do zrozumienia i interpretacji.

9.4. Wytyczne dotyczące projektowania wizualizacji

Skuteczne projektowanie wizualizacji polega na zwiększeniu zdolności odbiorcy do zrozumienia i interakcji z danymi. Wiąże się to z utrzymaniem starannej równowagi między elementami estetycznymi a funkcjonalnością, przy czym wybór koloru, czcionki i układu odgrywa kluczową rolę w jasnym i skutecznym przekazywaniu informacji. Prostota powinna być również zasadą przewodnią. Jednak według Cairo (2013) grafika nie powinna upraszczać wiadomości. Powinna je wyjaśniać, podkreślać trendy, odkrywać wzorce i ujawniać realia, które wcześniej nie były widoczne.

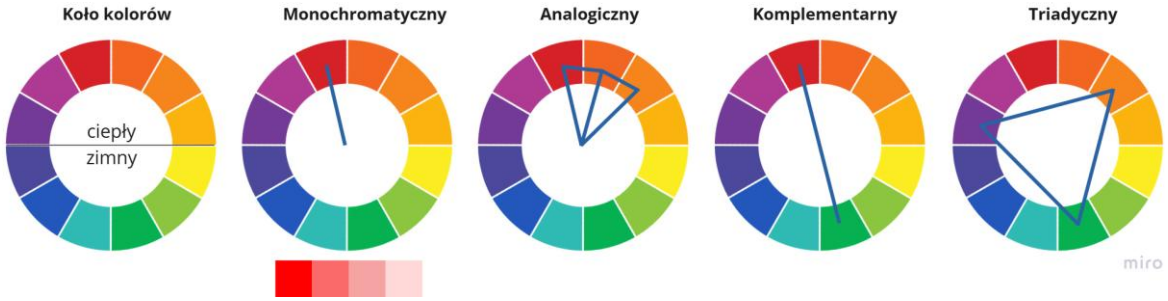
Częstą pułapką jest nadmierne komplikowanie wizualizacji zbyt wieloma elementami, które raczej dezorientują niż wyjaśniają. Celem jest uczynienie danych dostępnymi i zrozumiałymi dla docelowych odbiorców oraz zapewnienie, że wizualizacja służy swojemu celowi, którym jest informowanie i wspieranie podejmowania decyzji.

Przejrzystość i skuteczność wizualizacji danych można poprawić, usuwając niepotrzebne elementy rozpraszające uwagę. Nussbaumer Knaflic (2015) podkreśla, że każdy element, który nie dodaje wartości lub bezpośrednio nie wspiera zrozumienia danych, jest uważany za **bałagan**. Obejmuje to niepotrzebne linie siatki, nadmierne kolory, nieistotne punkty danych i zbyt złożone dekoracje wykresów. Elementy te mogą odwracać uwagę od kluczowych komunikatów, które dane mają przekazywać. Zaleca takie techniki, jak uproszczenie schematów kolorów, zminimalizowanie ilości tekstu i wykorzystanie białej przestrzeni. Dzięki strategicznemu wykorzystaniu białej przestrzeni można stworzyć wizualną hierarchię, która podkreśla kluczowe punkty danych i sprawia, że ogólna prezentacja jest bardziej przejrzysta i łatwiejsza do zrozumienia. Ponadto efektywne wykorzystanie białej przestrzeni może pomóc w stworzeniu zrównoważonego układu, który jest mniej zagracony i bardziej zorganizowany.

Kolor jest bardzo ważną częścią każdej wizualizacji danych. Służy nie tylko przyciąganiu uwagi, ale także organizowaniu informacji i skutecznemu przekazywaniu znaczenia. Prawidłowo użyty kolor może znacznie zwiększyć przejrzystość i wpływ wizualizacji. Istnieje kilka sugestii dotyczących skutecznego wykorzystania kolorów w wizualizacjach (Few, 2012; Cairo, 2013; Few, 2013; Nussbaumer Knaflic, 2015; Wexler i. in., 2017; Schwabish, 2021; Lidwell, 2023; Interaction Design Foundation, b.d.):

- **Wybierz odpowiednią kombinację kolorów:** korzystając z koła kolorów, można tworzyć wizualizacje, które są wizualnie zrównoważone i przyjemne dla oka. Istnieje kilka typowych kombinacji kolorów (Rysunek 9.17):
  - **Monochromatyczny** - jeden kolor w różnych odcieniach.
  - **Analogiczny -** trzy kolory znajdujące się obok siebie na kole kolorów. Kolory te są przyjemne dla oka i tworzą harmonijny projekt.
  - **Komplementarny** - dwa przeciwstawne kolory na kole kolorów. Kolory te są kontrastowe i powinny być używane do podkreślenia czegoś (np. wzrost - zielony/spadek - czerwony).
  - **Triadyczny** - trzy jednakowo odległe kolory na kole kolorów. Kolory te są dynamiczne i przyciągają uwagę**.**

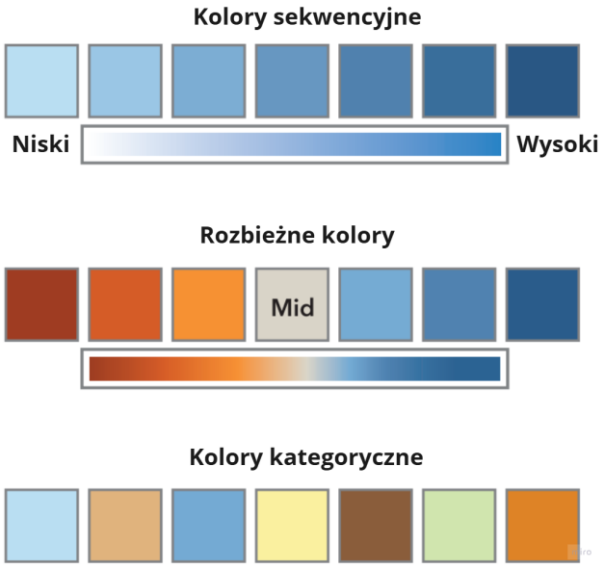
Ponadto, cieplejsze kolory powinny być używane dla elementów pierwszoplanowych, a chłodniejsze dla elementów tła.



Rysunek 9.17. Kombinacje kolorów

Źródło: Opracowanie własne na podstawie Lidwell (2023)

- **Wybierz odpowiednie schematy kolorów:** wybór schematu kolorów zależy od rodzaju danych, które mają być wizualizowane. **Sekwencyjny** schemat kolorów polega na użyciu tego samego koloru od jasnego do ciemnego odcienia. Jest on idealny do wyświetlania danych liczbowych, które przechodzą od niskiej do wysokiej wartości (np. sprzedaż według stanu). **Rozbieżny** schemat kolorów jest przydatny do podkreślania wartości powyżej lub poniżej punktu środkowego (np. zysku/straty). **Kategoryczny** schemat kolorów jest najlepszy dla danych kategorycznych, w których kolory muszą rozróżniać różne grupy bez sugerowania kolejności lub wartości (np. kategorie produktów). Rysunek 9.18 przedstawia różne schematy kolorów używane w wizualizacjach.



Rysunek 9.18. Schematy kolorów

Źródło: Wexler i in. (2017)

- **Używaj kolorów oszczędnie:** nadmierne użycie kolorów może powodować zamieszanie i utrudniać zrozumienie wykresu. Paleta kolorów powinna być ograniczona do tego, co ludzkie oko może szybko rozróżnić, czyli około pięciu różnych kolorów.
- **Rozważ ślepotę barw:** około 8% mężczyzn i 0,5% kobiet nie rozróżnia kolorów. Unikaj kombinacji kolorów, które są trudne do rozróżnienia dla użytkowników nierozróżniających kolorów, takich jak czerwony i zielony. Zamiast tych kolorów lepszym połączeniem jest pomarańczowy i niebieski.
- **Kolory powinny być spójne:** spójne użycie kolorów w wielu wizualizacjach pozwala odbiorcy łatwo zrozumieć i porównać dane. Po ustaleniu schematu kolorów dla określonych typów danych lub kategorii należy go zachować we wszystkich powiązanych wizualizacjach.
- **Użyj koloru, aby wyróżnić ważne dane:** kolor może być silnym wskaźnikiem tego, gdzie patrzeć. Użycie jasnego lub kontrastowego koloru może zwrócić uwagę na kluczowe punkty danych lub ustalenia, podczas gdy bardziej neutralne kolory mogą być używane do mniej krytycznych informacji. Niektórzy autorzy sugerują, że tworzenie jasnej, zrozumiałej wizualizacji powinno rozpocząć się od szarego koloru. Wszystkie elementy danych na wykresie (np. słupki na wykresie słupkowym lub linie na wykresie liniowym) powinny być szare. Następnie dodać należy etykiety i kolor tylko dla elementów, które chce się wyróżnić.

- **Zachowaj prostotę:** w dziedzinie wizualizacji danych zasada KISS, skrót od „Keep It Simple, Stupid", jest bardzo istotna i pomocna. Odnosi się ona do korzystania z prostych wykresów, ponieważ złożone wykresy lub zbyt szczegółowe wizualizacje mogą przytłoczyć użytkowników i utrudnić rozpoznanie kluczowych komunikatów lub punktów danych. Oznacza to również, że należy unikać wizualnego bałaganu i ograniczać niepotrzebne elementy wizualne, takie jak zbyt jasne kolory, czcionki i linie siatki. Stosując zasadę KISS, należy skupić się na samych danych, a nie na dekoracyjnych lub zbyt skomplikowanych elementach projektu.

Zasady te mają kluczowe znaczenie dla zapewnienia, że wizualizacje danych osiągają swój główny cel, jakim jest przekazywanie złożonych informacji w sposób dostępny i zrozumiały dla wszystkich odbiorców.

Z niniejszego rozdziału jasno wynika, że skuteczna wizualizacja danych jest kluczowym elementem w procesie podejmowania decyzji opartych na danych. W oparciu o zrozumienie kontekstu sytuacyjnego, w niniejszym rozdziale podkreślono znaczenie dostosowania wizualizacji do konkretnych potrzeb odbiorców docelowych. Poprzez szczegółową analizę różnych metod wizualizacji i zasad projektowania wykazano, że przemyślana wizualna reprezentacja danych jest ważna, aby umożliwić lepsze zrozumienie i komunikację złożonych informacji.

BIBLIOGRAFIA

1. Brush, K. (2022). Data visualization. TechTarget [dostępne: https://www.techtarget.com/searchbusinessanalytics/definition/data-visualization, dostęp 15.04.2024]
2. Cairo, A. (2013). The functional art: An introduction to information graphics and visualization. New Riders.
3. Data Visualisation Catalogue (b.d.). Scatterplot [dostępne: https://datavizcatalogue.com/methods/scatterplot.html, dostęp 17.04.2024]
4. Few, S. (2012). Show Me the Numbers. Analytics Press.
5. Few, S. (2013). Information dashboard design: Displaying data for at-a-glance monitoring. Analytics Press.
6. GeeksForGeeks (2024). What is Data Visualization and Why is It Important? dostępne: https://www.geeksforgeeks.org/data-visualization-and-its-importance/, dostęp 15.04.2024]
7. IBM (b.d.). What is data visualization? [dostępne: https://www.ibm.com/topics/data-visualization, dostęp 15.04.2024]
8. Interaction Design Foundation (b.d.). Color Theory [dostępne: https://www.interaction-design.org/literature/topics/color-theory, dostęp 18.04.2024]
9. Lidwell, W., Holden, K. & Butler, J. (2023). Universal Principles of Design, 3rd Edition. Quarto Publishing Group USA.
10. Nussbaumer Knaflic, C. (2015). Storytelling with data: A data visualization guide for business professionals. Wiley.
11. Schwabish (2020). Ten guidelines for better tables. Journal of Benefit-Cost Analysis, 11(2), pp. 151-178.
12. Schwabish (2021). Better data visualization: A guide for scholars, researchers and wonks. Columbia university press.
13. Tay, J. (2024). Effective use of BANs. Medium [dostępne: https://medium.com/@e0373084/eye-catching-bans-88d29632e4fa, dostęp 12.04.2024]
14. Wexler, S., Shaffer, J. & Cotgreave, A. (2017). The big book of dashboards: Visualizing your data using real-world business scenarios. Wiley.

15. Yi, M. (b.d.a). A complete guide to line charts. Attlassian [dostępne: https://www.atlassian.com/data/charts/line-chart-complete-guide, dostęp 17.04.2024]
16. Yi, M. (b.d.b). A complete guide to heatmaps. Attlassian [dostępne: https://www.atlassian.com/data/charts/heatmap-complete-guide, dostęp 17.04.2024]

10. ETYKA DANYCH I BEZPIECZEŃSTWO INFORMACJI

Autor: Dario Šebalj

W erze transformacji cyfrowej etyczne postępowanie z danymi i ich bezpieczeństwo stały się głównymi kwestiami dla osób fizycznych i organizacji. Ponieważ codziennie gromadzone i przetwarzane są ogromne ilości danych osobowych i wrażliwych, bardzo ważne jest zapewnienie odpowiedzialnego i bezpiecznego zarządzania tymi danymi.

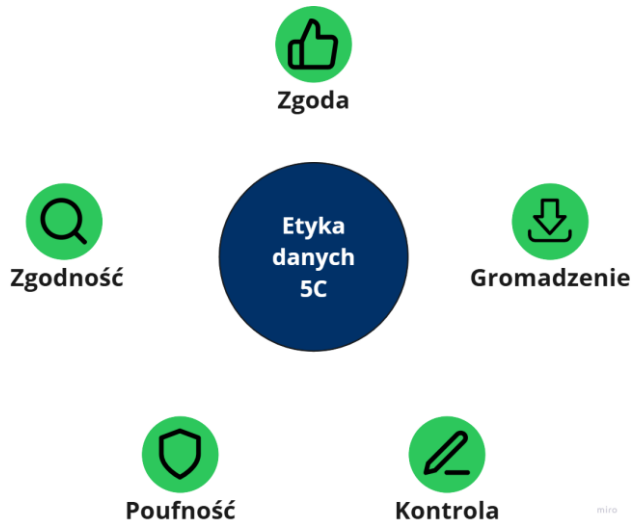
Niniejszy rozdział analizuje zasady etyki danych, podkreślając względy moralne i najlepsze praktyki w zakresie przetwarzania danych, a także bada różne zagrożenia dla bezpieczeństwa informacji. Rozumiejąc i zajmując się tymi kwestiami, możemy chronić prywatność, utrzymywać zaufanie i wspierać bezpieczniejsze środowisko cyfrowe.

10.1. Znaczenie etyki danych

Etyka danych odnosi się do zasad moralnych i praktyk kierujących gromadzeniem, przetwarzaniem, udostępnianiem i wykorzystywaniem danych w celu zapewnienia poszanowania praw osób fizycznych, dobrobytu społecznego i zaufania. Obejmuje przejrzystość, odpowiedzialność, uczciwość i prywatność, zapewniając, że praktyki dotyczące danych są zgodne ze standardami etycznymi i ramami prawnymi, celem zapobiegania szkodom i promowania odpowiedzialnych innowacji (Cognizant, b.d.; Gov.uk, 2020; Knight, 2021; McKinsey, 2022; Cepelak, 2023).

W dzisiejszym cyfrowym krajobrazie etyczne obchodzenie się z danymi ma kluczowe znaczenie dla utrzymania zaufania i zapewnienia przewagi konkurencyjnej. Artykuł McKinsey (2022) odnoszący się do tematu etyki danych podkreśla znaczenie włączenia kwestii etycznych do praktyk zarządzania danymi. Wskazuje na trzy powszechne błędy: zakładanie, że etyka danych jest nieistotna, poleganie wyłącznie na zespołach prawnych i zgodności w zakresie nadzoru oraz przedkładanie krótkoterminowych zysków finansowych nad praktyki etyczne. Aby rozwiązać te kwestie, zalecają kilka strategii. Po pierwsze, firmy powinny ustanowić jasne, specyficzne dla firmy wytyczne dotyczące etyki danych. Wytyczne te służą jako podstawa etycznego zarządzania danymi i pomagają w ustanowieniu standardu w całej organizacji.

Po drugie, tworzenie zróżnicowanych zespołów zajmujących się kwestiami związanymi z danymi zapewnia szeroki zakres perspektyw i zmniejsza ryzyko stronniczego podejmowania decyzji. Po trzecie, zaangażowanie kierownictwa wyższego szczebla jako mistrzów inicjatyw w zakresie etyki danych ma kluczowe znaczenie dla kierowania tymi praktykami w całej organizacji.



Rysunek 10.1. 5C Etyki Danych

Źródło: Opracowanie własne na podstawie Atlan (2023)

Rysunek 10.1 przedstawia 5C etyki danych, nakreślone przez Atlana (2023), które reprezentują podstawowe zasady etycznego postępowania z danymi:

- **Zgoda (ang. Consent):** uzyskanie świadomej, dobrowolnej zgody od osób fizycznych przed zebraniem ich danych, zapewnienie przejrzystości w zakresie wykorzystania danych.
- **Gromadzenie (ang. Collection):** gromadzenie tylko niezbędnych danych do określonych celów, unikając nadmiernego gromadzenia danych.
- **Kontrola (ang. Control):** umożliwienie osobom fizycznym dostępu do ich danych, ich przeglądania i aktualizowania, zapewniając im kontrolę nad ich wykorzystaniem.
- **Poufność (ang. Confidentiality):** ochrona danych przed nieautoryzowanym dostępem i naruszeniami dzięki solidnym środkom bezpieczeństwa.
- **Zgodność (ang. Compliance):** przestrzeganie wymogów prawnych i regulacyjnych, przeprowadzanie regularnych audytów w celu zapewnienia ciągłej zgodności.

Podobnie do zasad Atlana, Cote (2021) identyfikuje pięć podstawowych zasad etyki danych, które są niezbędne do przestrzegania przez profesjonalistów biznesowych:

- **Własność** podkreśla, że osoby fizyczne zachowują prawo własności do swoich danych osobowych. Gromadzenie danych osobowych bez wyraźnej zgody jest zarówno niezgodne z prawem, jak i nieetyczne. Firmy muszą uzyskać zgodę za pośrednictwem jasnych umów lub cyfrowych polityk prywatności, zapewniając, że użytkownicy są świadomi i zgadzają się na praktyki gromadzenia danych.
- **Przejrzystość** obejmuje jasną komunikację dotyczącą sposobu gromadzenia, przechowywania i wykorzystywania danych. Firmy muszą informować osoby fizyczne o metodach i celach gromadzenia danych. Taka przejrzystość buduje zaufanie i pozwala użytkownikom podejmować świadome decyzje dotyczące ich danych. Zwodnicze praktyki lub zatajanie informacji o wykorzystaniu danych są zarówno nieetyczne, jak i niezgodne z prawem.
- **Prywatność** koncentruje się na odpowiedzialności firm za ochronę prywatności danych osobowych. Nawet za zgodą, dane osobowe nie powinny być udostępniane publicznie bez wyraźnej zgody danej osoby. Firmy muszą wdrożyć solidne środki bezpieczeństwa w celu ochrony danych osobowych przed nieautoryzowanym dostępem lub naruszeniami.
- **Intencja** odnosi się do etycznych motywacji stojących za gromadzeniem i wykorzystywaniem danych. Dane powinny być gromadzone i wykorzystywane do celów, które są korzystne i nieszkodliwe dla jednostek lub społeczeństwa. Etyczne praktyki w zakresie danych obejmują wykorzystywanie danych w celu poprawy doświadczeń użytkowników i ulepszenia usług bez wykorzystywania lub powodowania szkód.
- **Wynik** uwzględnia szerszy wpływ wykorzystania danych na jednostki i społeczeństwo. Firmy muszą ocenić potencjalne konsekwencje swoich praktyk w zakresie danych i dążyć do uniknięcia negatywnych skutków. Zasada ta podkreśla potrzebę etycznego przewidywania i odpowiedzialności w podejmowaniu decyzji opartych na danych.

Guzman i Dyer (2020) podkreślają, że wyzwania etyczne w praktykach związanych z danymi nie są proste i często brakuje jednoznacznych rozwiązań. Stwierdzili, że istnieje rozbieżność między oczekiwaniami etycznymi online i offline. Wiele osób postrzega pewną formę wyjątkowości w przestrzeni online, gdzie tradycyjne zasady etyczne wydają się nie mieć zastosowania. Taki sposób myślenia może prowadzić do usprawiedliwiania działań online, które zostałyby uznane za nieetyczne offline. Autorzy proponują podejście etyczne, które łączy obie sfery, podkreślając, że zasady etyczne powinny pozostać spójne niezależnie od medium.

Artykuł na temat etyki danych opublikowany przez Basl i in. (2021) bada złożony proces przechodzenia od abstrakcyjnych zasad etycznych do konkretnych, możliwych do wdrożenia obietnic w kontekście dużych zbiorów danych i sztucznej inteligencji (AI). Doszli do wniosku, że dokonanie tej zmiany jest trudne i ważne, aby zagwarantować, że etyczne zachowanie nie jest tylko teoretyczne, ale także praktyczne i znaczące.

Cztery anonimowe studia przypadków opracowane przez Princeton's Center for Information Technology Policy i Center for Human Values w celu zachęcenia do dyskursu etycznego zostały opisane przez O'Reilly'ego (2018). Jedno ze studiów przypadku bada dylematy etyczne, jakie stwarza zautomatyzowana aplikacja opieki zdrowotnej zaprojektowana w celu pomocy pacjentom z cukrzycą u dorosłych przy użyciu sztucznej inteligencji. Podkreśla ono potrzebę zrównoważenia korzyści technologicznych z zasadami etycznymi, takimi jak autonomia, sprawiedliwość i odpowiedzialność. Podjęcie tych wyzwań etycznych ma kluczowe znaczenie dla odpowiedzialnej integracji sztucznej inteligencji w opiece zdrowotnej, zapewniając, że służy ona najlepszym interesom wszystkich pacjentów. Istnieje kilka kluczowych kwestii, którymi należy się zająć:

- **Paternalizm:** aplikacja ma na celu zachęcanie pacjentów do zdrowszych zachowań poprzez nakłanianie ich do dokonywania lepszych wyborów. Chociaż takie działanie może poprawić wyniki zdrowotne, rodzi to pytania etyczne dotyczące autonomii i paternalizmu. Czy etyczne jest wpływanie przez aplikację na zachowania pacjentów, czy też powinni oni mieć pełną autonomię w podejmowaniu decyzji zdrowotnych?
- **Zgoda i przejrzystość:** Aplikacja gromadzi wrażliwe dane zdrowotne w celu skutecznego działania. Zapewnienie świadomej zgody i przejrzystości w zakresie gromadzenia, wykorzystywania i udostępniania danych ma kluczowe znaczenie. Pacjenci muszą być w pełni świadomi tego, jakie dane są gromadzone, w jaki sposób będą wykorzystywane i kto będzie miał do nich dostęp.
- **Prywatność i bezpieczeństwo danych:** Obsługa wrażliwych danych medycznych wymaga rygorystycznych środków ochrony prywatności i bezpieczeństwa. Studium przypadku podkreśla potrzebę solidnych protokołów ochrony danych w celu zabezpieczenia informacji o pacjentach przed naruszeniami i nieautoryzowanym dostępem.
- **Odpowiedzialność i rozliczalność:** Ustalenie, kto jest odpowiedzialny za decyzje i działania aplikacji jest kolejnym kluczowym aspektem. Jeśli aplikacja wyda nieprawidłowe zalecenie, które negatywnie wpłynie na zdrowie pacjenta, identyfikacja

podmiotu odpowiedzialnego (deweloperów, świadczeniodawców opieki zdrowotnej lub samej aplikacji) jest złożona, ale niezbędna do pociągnięcia do odpowiedzialności.

Nowoczesne zarządzanie danymi opiera się zasadniczo na etyce danych, która gwarantuje uczciwe, przejrzyste, odpowiedzialne i szanujące prywatność praktyki w zakresie danych. Organizacje mogą wspierać odpowiedzialne innowacje, unikać pułapek prawnych i zwiększać zaufanie poprzez przestrzeganie standardów etycznych. Włączenie silnych ram etycznych do praktyk zarządzania danymi jest nie tylko najlepszą praktyką, ale także wymogiem zrównoważonego i odpowiedzialnego rozwoju, ponieważ dane stają się coraz bardziej istotne dla operacji i podejmowania decyzji.

Innym ważnym aspektem jest bezpieczeństwo informacji, ponieważ etyka danych i bezpieczeństwo informacji są ze sobą nierozerwalnie związane. Zapewnienie etycznych praktyk w zakresie danych stanowi podstawę dla solidnych środków bezpieczeństwa informacji. Ochrona danych przed nieautoryzowanym dostępem, naruszeniami i innymi zagrożeniami bezpieczeństwa nie tylko chroni prywatność i poufność, ale także podtrzymuje zasady etyczne omówione w tym rozdziale.

10.2. Podstawy bezpieczeństwa informacji

Bezpieczeństwo informacji odnosi się do kompleksowego zestawu praktyk i zasad mających na celu ochronę informacji i systemów informatycznych przed nieautoryzowanym dostępem, wykorzystaniem, ujawnieniem, zakłóceniem, modyfikacją lub zniszczeniem. Zapewnia poufność, integralność i dostępność danych poprzez wdrożenie środków ochronnych, polityk i technologii. Środki te obejmują kontrolę dostępu, szyfrowanie, odzyskiwanie po awarii oraz zgodność z normami prawnymi i regulacyjnymi w celu ograniczenia ryzyka i ochrony przed potencjalnymi zagrożeniami (Fruhlinger, 2020; CISCO, b.d., NIST, b.d.).

Bezpieczeństwo informacji ma obecnie zasadnicze znaczenie dla wiarygodności i integralności organizacji w erze cyfrowej. Jego znaczenie podkreśla rosnąca zależność od danych cyfrowych i wzrost cyberzagrożeń, które narażają wrażliwe dane. Przede wszystkim bezpieczeństwo informacji chroni wrażliwe dane przed nieautoryzowanym dostępem, naruszeniami i kradzieżą. Obejmuje to dane osobowe, dane finansowe, własność intelektualną i poufną komunikację biznesową. W miarę jak cyberataki stają się coraz bardziej wyrafinowane, ryzyko naruszenia danych rośnie, potencjalnie prowadząc do poważnych strat finansowych

i utraty reputacji. Na przykład naruszenie Equifax z 2017 r. ujawniło dane osobowe 147 milionów osób, co skutkowało ugodą w wysokości do 425 milionów dolarów (Federalna Komisja Handlu, 2022). Takie incydenty podkreślają tragiczne konsekwencje nieodpowiednich środków bezpieczeństwa informacji.

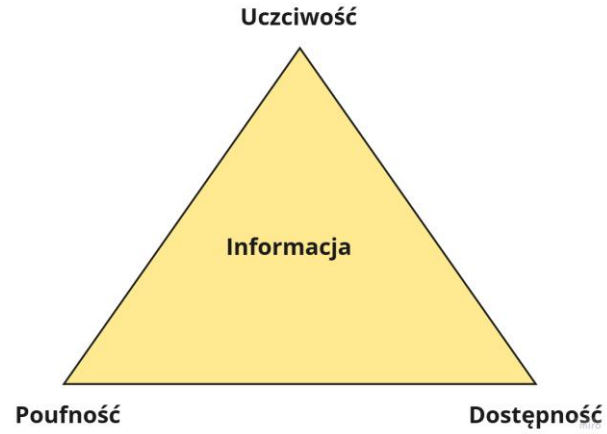
Co więcej, bezpieczeństwo informacji ma kluczowe znaczenie dla utrzymania zaufania konsumentów. W czasach, gdy prywatność danych ma kluczowe znaczenie, klienci są coraz bardziej zaniepokojeni tym, w jaki sposób przetwarzane są ich informacje. Silna architektura bezpieczeństwa informacji zapewnia, że dane konsumentów są bezpieczne, co sprzyja lojalności i zaufaniu. Według ankiety IBM, **75%** klientów nie dokonałoby zakupów w firmie, której nie ufają co do ochrony danych (PR Newswire, 2018). Bezpieczeństwo informacji jest zatem zarówno potrzebą technologiczną, jak i strategicznym imperatywem biznesowym.

Bezpieczeństwo informacji ma również kluczowe znaczenie dla łagodzenia zakłóceń operacyjnych. Cyberataki, takie jak oprogramowanie ransomware, mogą zakłócać operacje korporacyjne, uniemożliwiając użytkownikom dostęp do niezbędnych systemów do czasu zapłacenia okupu. Atak ransomware na Colonial Pipeline w 2021 r., który doprowadził do niedoborów paliwa we wschodnich Stanach Zjednoczonych, stanowi przykład destrukcyjnego potencjału takich zagrożeń (Kerner, 2022). Wdrażając silne środki bezpieczeństwa, organizacje mogą zabezpieczyć swoją ciągłość operacyjną i odporność na takie zakłócenia.

Według Kim i Solomon (2018) informacje są uważane za bezpieczne, jeśli spełniają trzy główne założenia:

- **Poufność (ang. Confidentiality):** dostęp do poufnych informacji mają tylko upoważnione osoby.
- **Uczciwość (ang. Integrity):** informacje mogą być zmieniane tylko przez osoby posiadające odpowiednie uprawnienia.
- **Dostępność (ang. Availability):** informacje i zasoby są dostępne dla autoryzowanych użytkowników zawsze, gdy są potrzebne.

Te założenia są często nazywane triadą CIA, jak pokazano na Rysunku 10.2.



Rysunek 10.2. Triada CIA

Źródło: Opracowane własne na podstawie Kim i Solomon (2018)

W bezpieczeństwie informacji pojęcia ryzyka, zagrożenia i podatności są kluczowe dla zrozumienia i zarządzania bezpieczeństwem. **Ryzyko** to prawdopodobieństwo, że coś złego stanie się z zasobem. W bezpieczeństwie informacji ryzyko to prawdopodobieństwo wystąpienia szkodliwego zdarzenia wpływającego na poufność, integralność lub dostępność informacji. Na przykład firma może ocenić ryzyko cyberataku na swoją sieć, biorąc pod uwagę zarówno prawdopodobieństwo takiego ataku, jak i potencjalne szkody, jakie może on spowodować. **Zagrożenie** to każde działanie, które może potencjalnie wyrządzić szkodę systemom informatycznym lub danym. Zagrożenia mogą być celowe, takie jak cyberataki hakerów, lub niezamierzone, takie jak klęski żywiołowe lub błędy ludzkie. **Podatność** odnosi się do słabości lub luk w zabezpieczeniach systemu, które mogą zostać wykorzystane przez zagrożenia w celu wyrządzenia szkody. Mogą to być wady oprogramowania, sprzętu, procesów organizacyjnych lub ludzkich zachowań (Kim i Solomon, 2018).

Aby skutecznie chronić systemy informatyczne, organizacje muszą zrozumieć, w jaki sposób ryzyko, zagrożenia i luki w zabezpieczeniach wzajemnie na siebie oddziałują. Na przykład luka w oprogramowaniu (np. luka w zabezpieczeniach) może zostać wykorzystana przez podmiot stanowiący zagrożenie (np. hakera), prowadząc do naruszenia danych, co stanowi ryzyko dla organizacji. Identyfikując i ograniczając luki w zabezpieczeniach, organizacje mogą zmniejszyć ryzyko wystąpienia zagrożeń powodujących znaczne szkody.

10.2.1. Naruszenie danych

Naruszenia danych stały się poważnym zagrożeniem w erze cyfrowej, wpływając na organizacje z różnych sektorów. Naruszenia te narażają na szwank poufne informacje,

prowadząc do strat finansowych, utraty reputacji i konsekwencji prawnych. Zrozumienie skali i wpływu naruszeń danych ma kluczowe znaczenie dla opracowania skutecznych strategii bezpieczeństwa w celu ochrony przed takimi incydentami. Według Fortinet (b.d.), **naruszenie danych** „to zdarzenie, które skutkuje ujawnieniem poufnych, prywatnych, chronionych lub wrażliwych informacji osobie nieupoważnionej do dostępu do nich". Naruszenia te mogą wystąpić na różne sposoby (Kaspersky, b.d.a):

1. **Przypadkowy insider (osoba mająca dostęp do poufnych informacji):** Pracownik nieumyślnie uzyskuje dostęp do poufnych informacji bez odpowiedniego upoważnienia. Na przykład, korzystając z komputera współpracownika i przeglądając poufne pliki.
2. **Złośliwy insider:** Osoba z autoryzowanym dostępem celowo nadużywa danych w szkodliwych celach. Może to obejmować kradzież lub wyciek poufnych informacji.
3. **Zgubione lub skradzione urządzenia:** Niezaszyfrowane i niezabezpieczone urządzenia, takie jak laptopy lub dyski zewnętrzne zawierające poufne dane, zostają zgubione lub skradzione, przez co informacje są narażone na nieautoryzowany dostęp.
4. **Złośliwi przestępcy zewnętrzni:** Hakerzy wykorzystują różne metody do włamywania się do systemów, w tym ataki phishingowe, ataki siłowe i złośliwe oprogramowanie. Cyberprzestępcy wykorzystują luki w oprogramowaniu, sieciach i zachowaniu użytkowników, aby uzyskać dostęp do wrażliwych danych.

Według Kaspersky (n.d.a), powszechne metody wykorzystywane w naruszeniach danych obejmują **phishing**, w którym cyberprzestępcy podszywają się pod zaufane podmioty, aby nakłonić osoby do ujawnienia poufnych informacji. Inną metodą są **ataki siłowe (brute force)**, w których hakerzy używają oprogramowania do wielokrotnego odgadywania haseł, wykorzystując słabe lub wielokrotnie używane dane uwierzytelniające w celu uzyskania nieautoryzowanego dostępu do kont. Ponadto **złośliwe oprogramowanie (malware)**, takie jak oprogramowanie szpiegujące, jest wykorzystywane do infiltracji systemów i kradzieży danych w sposób niezauważony. Metody te zostaną wyjaśnione w dalszej części rozdziału.

Błąd ludzki odpowiada za **95%** wszystkich naruszeń danych (Cybernews, 2022).

XXI wiek był świadkiem jednych z najpoważniejszych naruszeń danych, podkreślających luki w systemach cyfrowych i krytyczną potrzebę stosowania solidnych

środków bezpieczeństwa. Według Hill i Swinhoe (2022) oraz ESET (n.d.), wśród niektórych z najbardziej znaczących naruszeń danych wyróżnić można:

- **Yahoo (2013-2014):** Yahoo doświadczyło jednego z największych naruszeń danych w historii - w 2013 roku włamano się do wszystkich trzech miliardów kont użytkowników. Naruszenie to ujawniło nazwiska, adresy e-mail, daty urodzenia oraz pytania i odpowiedzi zabezpieczające. Kolejne naruszenie w 2014 r. dotknęło 500 milionów kont, jeszcze bardziej podkreślając luki w zabezpieczeniach firmy.
- **Marriott International (2018):** Marriott ogłosił, że doszło do naruszenia danych około 500 milionów gości. Hakerzy uzyskali dostęp do bazy danych rezerwacji gości Starwood, ujawniając dane osobowe, takie jak nazwiska, adresy, numery telefonów, adresy e-mail i numery paszportów. Naruszenie to było wynikiem nieautoryzowanego dostępu sięgającego 2014 roku. Biuro Komisarza ds. Informacji (ICO), brytyjski organ regulacyjny ds. danych, ostatecznie ukarał korporację grzywną w wysokości 18,4 miliona funtów w 2020 roku za brak ochrony prywatności danych osobowych swoich klientów.
- **Adult Friend Finder (2016):** Naruszenie Adult Friend Finder ujawniło dane osobowe 412 milionów kont, w tym imiona i nazwiska, adresy e-mail i hasła, z których wiele było słabo zaszyfrowanych. Incydent ten wzbudził poważne obawy dotyczące praktyk bezpieczeństwa platform internetowych obsługujących wrażliwe dane osobowe.
- **MySpace (2013):** Naruszenie danych MySpace, które miało miejsce w 2013 roku, spowodowało ujawnienie ponad 360 milionów kont użytkowników. Dane obejmowały imiona i nazwiska, adresy e-mail i hasła. Hakerzy sprzedali później te informacje w ukrytej sieci (ang. dark web), co uwypukliło luki w systemach bezpieczeństwa platform mediów społecznościowych w tym czasie.
- **LinkedIn (2021):** W 2021 roku dane osobowe 700 milionów użytkowników LinkedIn zostały opublikowane na forum dark web. Haker wykorzystał techniki screen scrapingu (wydobywania danych) za pośrednictwem interfejsu API LinkedIn w celu uzyskania adresów e-mail, numerów telefonów i innych danych osobowych. Chociaż nie był to tradycyjny atak hakerski, incydent ten wzbudził poważne obawy dotyczące prywatności danych i potencjalnego niewłaściwego wykorzystania screen scrapingu do ataków socjotechnicznych.
- **Equifax (2017):** Naruszenie to ujawniło dane osobowe prawie 148 milionów Amerykanów, 15,2 miliona Brytyjczyków i 19 000 Kanadyjczyków. Hakerzy wykorzystali

lukę we frameworku aplikacji internetowej Apache Struts, której Equifax nie zdołał załatać. Skradzione dane obejmowały numery ubezpieczenia społecznego, daty urodzenia i adresy, co doprowadziło do kosztów szacowanych na 1,7 miliarda dolarów dla Equifax.

- **eBay (2014):** eBay ujawnił naruszenie, które dotknęło 145 milionów użytkowników. Atak został przeprowadzony z wykorzystaniem naruszonych danych logowania pracowników, co doprowadziło do ujawnienia nazwisk, adresów e-mail, adresów fizycznych, numerów telefonów i zaszyfrowanych haseł. Naruszenie to uwypukliło luki w kontroli dostępu pracowników i znaczenie silnych środków uwierzytelniania.
- **Target (2013):** Naruszenie dotyczyło ponad 41 milionów kont kart płatniczych i danych kontaktowych ponad 60 milionów klientów. Cyberprzestępcy uzyskali dostęp do danych klientów, w tym nazwisk, numerów telefonów, adresów e-mail, numerów kart kredytowych i debetowych oraz zaszyfrowanych kodów PIN. Target poniósł znaczne koszty prawne i ugodowe, w tym pozew zbiorowy o wartości 10 milionów dolarów i ugodę wielostanową o wartości 18,5 miliona dolarów.

Naruszenia te pokazują daleko idące konsekwencje cyberataków i krytyczną potrzebę stosowania silnych praktyk w zakresie cyberbezpieczeństwa. Ucząc się na podstawie tych głośnych przypadków, organizacje mogą lepiej chronić swoje dane, ulepszać swoje protokoły bezpieczeństwa i minimalizować ryzyko przyszłych naruszeń. Naruszenia danych są często wynikiem szeregu podstawowych zagrożeń. Zrozumienie tych zagrożeń ma kluczowe znaczenie dla budowania skutecznych środków bezpieczeństwa informacji. Zagrożenia cybernetyczne mogą pochodzić z różnych źródeł, w tym od złośliwych intruzów, cyberprzestępców, a nawet podmiotów sponsorowanych przez państwo. Ponadto luki w systemach i sieciach mogą zostać wykorzystane do uzyskania nieautoryzowanego dostępu do poufnych informacji.

10.2.2. Zagrożenia dotyczące bezpieczeństwa informacji

Zagrożenie bezpieczeństwa to złośliwe działanie, które próbuje uszkodzić lub ukraść dane, zagrozić systemom organizacji lub całej firmie (TechTarget, 2024). Istnieje wiele różnych zagrożeń bezpieczeństwa informacji, które poważnie zagrażają dostępności, integralności i poufności danych.

Według Kim i Solomona (2018) oraz Grubba (2021) **złośliwe oprogramowanie (malware)** ma na celu infiltrację, uszkodzenie lub wyłączenie komputerów i sieci. Typowe rodzaje złośliwego oprogramowania obejmują wirusy, robaki, trojany, oprogramowanie

ransomware i oprogramowanie szpiegujące. **Wirus** to rodzaj złośliwego oprogramowania, które dołącza się do legalnego programu lub pliku i rozprzestrzenia się na inne programy i pliki po uruchomieniu zainfekowanego oprogramowania. Wirusy mogą uszkadzać lub usuwać dane, zakłócać działanie systemu i rozprzestrzeniać się na inne systemy za pośrednictwem załączników do wiadomości e-mail, połączeń sieciowych lub nośników wymiennych. W przeciwieństwie do wirusów, **robaki** są samodzielnym złośliwym oprogramowaniem, które może samoczynnie się powielać i rozprzestrzeniać niezależnie w sieciach bez konieczności dołączania do programu hosta. Robaki wykorzystują luki w systemach operacyjnych lub aplikacjach do rozprzestrzeniania się, często powodując przeciążenie sieci i przeciążenie systemów poprzez zużywanie przepustowości i zasobów. **Trojan** lub **koń trojański** to złośliwe oprogramowanie podszywające się pod legalne oprogramowanie. Użytkownicy są nakłaniani do jego instalacji, wierząc, że jest to nieszkodliwy lub przydatny program. **Ransomware** to rodzaj złośliwego oprogramowania, które szyfruje dane ofiary, czyniąc je niedostępnymi do czasu zapłacenia okupu atakującemu. Ataki ransomware mogą być niszczycielskie, prowadząc do znacznej utraty danych i zakłóceń operacyjnych, jeśli okup nie zostanie zapłacony lub kopie zapasowe nie będą dostępne. **Oprogramowanie szpiegujące** to złośliwe oprogramowanie zaprojektowane w celu gromadzenia informacji o osobie lub organizacji bez ich wiedzy Może gromadzić różne rodzaje danych, takie jak naciśnięcia klawiszy, nawyki przeglądania i dane osobowe, i przekazywać je stronom trzecim.

Innym rodzajem zagrożenia jest **phishing**. Kosinski (2024) wyjaśnia, że ataki phishingowe obejmują fałszywe wiadomości e-mail, SMS-y, połączenia telefoniczne lub strony internetowe mające na celu nakłonienie osób do ujawnienia danych osobowych lub pobrania złośliwego oprogramowania. Ataki te wykorzystują ludzkie błędy i zaufanie, dzięki czemu są bardzo skuteczne. Aby zwalczać phishing, organizacje muszą korzystać z zaawansowanych narzędzi do wykrywania zagrożeń i zapewniać solidne szkolenia pracowników, celem skutecznego rozpoznawania i reagowania na te oszustwa.

Phishing jest główną przyczyną naruszeń danych, stanowiąc 16% wszystkich naruszeń i kosztując organizacje średnio **4,76 miliona** dolarów za każde naruszenie (Kosinski, 2024).

Istnieją cztery główne rodzaje phishingu (Forbes, 2024):

- **Phishing mailowy:** wykorzystywanie poczty elektronicznej do kradzieży poufnych informacji. Atakujący mogą atakować duże grupy odbiorców, przyjmując tożsamość renomowanych organizacji.
- **Spear phishing:** wysyłanie zindywidualizowanych wiadomości e-mail, SMS-ów lub połączeń telefonicznych z zamiarem uzyskania dostępu do systemów komputerowych lub poufnych informacji. Korzystając z tej techniki, atakujący zazwyczaj wykorzystują dane z otwartych baz danych, mediów społecznościowych lub wcześniejszych naruszeń, aby wzmocnić swoją legalność.
- **Whaling:** koncentruje się na pracownikach wysokiego szczebla, w tym dyrektorach finansowych i dyrektorach naczelnych. Atakujący tworzą bardzo przekonującą, wysoce dostosowaną komunikację w celu uzyskania poufnych danych i informacji od firmy.
- **Vishing:** wykonywanie połączeń telefonicznych lub zostawianie wiadomości głosowych pod pozorem wiarygodnego źródła. Celem jest zdobycie kont bankowych, wykorzystanie danych osobowych i kradzież pieniędzy.

Zagrożenia wewnętrzne to zagrożenia bezpieczeństwa pochodzące z wewnątrz organizacji. Mogą to być pracownicy, kontrahenci lub partnerzy biznesowi, którzy mają dostęp do systemów i danych organizacji. Zagrożenia te mogą być szczególnie niebezpieczne, ponieważ insiderzy często mają legalny dostęp do poufnych informacji i systemów, co utrudnia wykrycie ich złośliwych działań (TechTarget, 2024).

Innym rodzajem zagrożenia są ataki **DDoS (Distributed Denial-of-Service)**. Ich celem jest zakłócenie normalnego ruchu docelowego serwera, usługi lub sieci poprzez przeciążenie go nadmiernym ruchem internetowym. Osiąga się to poprzez wykorzystanie wielu zainfekowanych systemów komputerowych jako źródeł ruchu atakującego. Gdy urządzenia te, często rozproszone globalnie, jednocześnie wysyłają liczne żądania do celu, zużywają dostępną przepustowość i zasoby, prowadząc do przerw w świadczeniu usług i uniemożliwiając legalnym użytkownikom dostęp do usługi (TechTarget, 2024).

Zagrożenia bezpieczeństwa w Internecie są ściśle związane z działaniami hakerów, którzy wykorzystują luki w systemach do różnych złośliwych celów. Według Grubba (2021) hakerzy są często kategoryzowani na podstawie ich intencji i metod. Dwie podstawowe kategorie to hakerzy w białych kapeluszach i hakerzy w czarnych kapeluszach. **Hakerzy w białych kapeluszach**, znani również jako etyczni hakerzy, wykorzystują swoje umiejętności do celów obronnych. Ich zadaniem jest ochrona organizacji przed cyberzagrożeniami poprzez identyfikowanie i naprawianie luk w zabezpieczeniach, zanim

złośliwi hakerzy będą mogli je wykorzystać. Z kolei **hakerzy w czarnych kapeluszach** angażują się w nielegalne działania o złych intencjach. Wykorzystują luki w zabezpieczeniach dla osobistych korzyści, które mogą obejmować kradzież danych, rozprzestrzenianie złośliwego oprogramowania lub powodowanie zakłóceń.

Jedną z kluczowych metod ochrony przed zagrożeniami bezpieczeństwa informacji jest stosowanie silnych haseł. Silne hasła, które powinny być złożone i unikalne dla każdego konta, znacznie zmniejszają ryzyko nieautoryzowanego dostępu.

Tabela 10.1 przedstawia czas potrzebny hakerowi na brutalne wymuszenie hasła, zgodnie z badaniami przeprowadzonymi przez Hive Systems (2024).

Tabela 10.1. Czas potrzebny hakerowi na złamanie hasła w 2024 r.

Liczba znaków	Tylko cyfry	Małe litery	Wielkie i małe litery	Cyfry, duże i małe litery	Cyfry, małe i wielkie litery, symbole
4	Natychmiast	Natychmiast	3 sekundy	6 sekund	9 sekund
5	Natychmiast	4 sekundy	2 minuty	6 minut	10 minut
6	Natychmiast	2 minuty	2 godziny	6 godzin	12 godzin
7	4 sekundy	50 minut	4 dni	2 tygodnie	1 miesiąc
8	37 sekund	22 godziny	8 miesięcy	3 lata	7 lat
9	6 minut	3 tygodnie	33 lat	161 lat	479 lat
10	1 godzina	2 lat	1k lat	9k lat	33k lat
11	10 godzin	44 lat	89k lat	618k lat	2m lat
12	4 dni	1k lat	4m lat	38m lat	164m lat
13	1 miesiąc	29k lat	241m lat	2bn lat	11bn lat
14	1 rok	766k lat	12bn lat	147bn lat	805bn lat
15	12 lat	19m lat	652bn lat	9tn lat	56tn lat
16	119 lat	517m lat	33tn lat	566tn lat	3qd lat
17	1k lat	13bn lat	1qd lat	35qd lat	276qd lat
18	11k lat	350bn lat	91qd lat	2qn lat	19qn lat

Źródło: Opracowanie własne na podstawie Hive Systems (2014)

Zrozumienie różnych rodzajów zagrożeń bezpieczeństwa informacji, takich jak ataki phishingowe, złośliwe oprogramowanie i ataki DDoS, podkreśla krytyczną potrzebę stosowania solidnych środków cyberbezpieczeństwa. Zagrożenia te stanowią poważne ryzyko dla danych osobowych, informacji finansowych i integralności organizacyjnej. Z tego powodu niezbędne staje się przyjęcie kompleksowych strategii bezpieczeństwa. Poniższy podrozdział przedstawia sugestie dotyczące utrzymania silnego bezpieczeństwa w Internecie, w tym praktyczne wskazówki, które ludzie i instytucje mogą wykorzystać do ochrony swoich zasobów cyfrowych.

10.2.3. Zalecenia dotyczące bezpieczeństwa informacji

Dużą liczbę zagrożeń bezpieczeństwa, takich jak phishing i złośliwe oprogramowanie, można znacznie zminimalizować, rozumiejąc je i stosując w praktyce. Istnieje kilka najważniejszych zaleceń dotyczących zapewnienia bezpieczeństwa informacji (Rubenking & Duffy, 2023; NSW Government, b.d.; Kaspersky, b.d.b):

- **Używaj silnych haseł:** twórz złożone hasła składające się z liter, cyfr i symboli dla każdego konta. Unikaj używania łatwych do odgadnięcia informacji, takich jak data urodzin. Używaj menedżera haseł do bezpiecznego przechowywania haseł i zarządzania nimi.
- Jeśli to możliwe, **włącz uwierzytelnianie wieloskładnikowe (MFA):** dodaj dodatkową warstwę zabezpieczeń, wymagając dwóch lub więcej metod weryfikacji w celu uzyskania dostępu do kont, takich jak hasło i jednorazowy kod wysłany na telefon.
- **Aktualizuj oprogramowanie:** regularnie aktualizuj systemy operacyjne, przeglądarki i aplikacje w celu załatania luk w zabezpieczeniach. W miarę możliwości włączaj automatyczne aktualizacje, aby mieć pewność, że zawsze jesteś chroniony przed najnowszymi zagrożeniami.
- **Bądź świadomy oszustw phishingowych:** nie klikaj linków ani nie pobieraj załączników z nieznanych lub podejrzanych wiadomości e-mail. Weryfikuj informacje o nadawcy i szukaj oznak phishingu, takich jak błędy ortograficzne lub pilne prośby o podanie danych osobowych.
- **Korzystaj z bezpiecznych połączeń:** upewnij się, że twoje połączenie internetowe jest bezpieczne, korzystając z wirtualnych sieci prywatnych (VPN) i unikając publicznych sieci Wi-Fi w przypadku wrażliwych działań, takich jak bankowość internetowa. Sprawdź „https://" w adresie URL, wskazujące na bezpieczne połączenie.
- **Regularne twórz kopie zapasowych danych:** regularnie twórz kopie zapasowe danych na dyskach zewnętrznych lub w chmurze. Ta praktyka zapewnia możliwość odzyskania informacji w przypadku awarii sprzętu, kradzieży lub ataku ransomware.
- **Zainstaluj oprogramowanie antywirusowe:** używaj renomowanego oprogramowania zabezpieczającego do wykrywania, zapobiegania i usuwania złośliwego oprogramowania. Aktualizuj oprogramowanie antywirusowe i wykonuj regularne skanowanie, aby upewnić się, że system jest czysty.

- **Regularne monitoruj konta:** często sprawdzaj swoje konta finansowe i internetowe pod kątem wszelkich nieautoryzowanych działań. Skonfiguruj alerty dla nietypowych transakcji i natychmiast zgłaszaj dostawcy usług wszelkie podejrzane zachowania.

Rozumiejąc etyczne implikacje przetwarzania danych i rozpoznając istniejące zagrożenia bezpieczeństwa, osoby i organizacje mogą opracować skuteczne strategie ochrony poufnych informacji. Od ustanowienia wytycznych etycznych i używania silnych, unikalnych haseł po wdrażanie zaawansowanych środków bezpieczeństwa i bycie na bieżąco z potencjalnymi zagrożeniami, praktyki te wspólnie zapewniają integralność, poufność i dostępność danych. Nadając priorytet etyce danych i solidnym protokołom bezpieczeństwa, można stworzyć bezpieczniejsze, bardziej godne zaufania środowisko cyfrowe.

BIBLIOGRAFIA

1. Atlan (2023). Data Ethics Unveiled: Principles & Frameworks Explored [dostępne: https://atlan.com/data-ethics-101/, dostęp 17.05.2024]
2. Basl, J., Sandler, R. & Tiell, S. (2021). Getting from commitment to content in AI and data ethics: Justice and explainability. Atlantic Council [dostępne: https://www.atlanticcouncil.org/in-depth-research-reports/report/specifying-normative-content/, dostęp 17.05.2024]
3. Cepelak, C. (2023). What is Data Ethics? Datacamp [dostępne: https://www.datacamp.com/blog/introduction-to-data-ethics, dostęp 14.05.2024]
4. CISCO (n.d.). What Is Information Security? [dostępne: https://www.cisco.com/c/en/us/products/security/what-is-information-security-infosec.html, dostęp 20.05.2024]
5. Cognizant (n.d.). Data ethics [dostępne: https://www.cognizant.com/us/en/glossary/data-ethics, dostęp 14.05.2024]
6. Cote (2021). 5 Princples of Data Ethics for Business. Harvard Business School Online [dostępne: https://online.hbs.edu/blog/post/data-ethics, dostęp 17.05.2024]
7. Cybernews (2022). World Economic Forum finds that 95% of cybersecurity incidents occur due to human error [dostępne: https://cybernews.com/editorial/world-economic-forum-finds-that-95-of-cybersecurity-incidents-occur-due-to-human-error/, dostęp 21.05.2024]
8. ESET (n.d.). 5 scary data breaches that shook the world [dostępne: https://www.eset.com/in/about/newsroom/corporate-blog/corporate-blog/eset-5-scary-data-breaches-that-shook-the-world/, dostęp 21.05.2024]
9. Federal Trade Commission (2022). Equifax Data Breach Settlement [dostępne: https://www.ftc.gov/enforcement/refunds/equifax-data-breach-settlement, dostęp 20.05.2024]
10. Forbes (2024). Cybersecurity Stats: Facts And Figures You Should Know dostępne: https://www.forbes.com/advisor/education/it-and-tech/cybersecurity-statistics/, dostęp 24.05.2024]
11. Fortinet (n.d.). What Is A Data Breach? [dostępne: https://www.fortinet.com/resources/cyberglossary/data-breach, dostęp 21.05.2024]

12. Fruhlinger, J. (2020). What is information security? Definition, principles, and jobs. CSO [dostępne: https://www.csoonline.com/article/568841/what-is-information-security-definition-principles-and-jobs.html, dostęp 20.05.2024]
13. Gov.uk (2020). Data Ethics Framework: glossary and methodology [dostępne: https://www.gov.uk/government/publications/data-ethics-framework/data-ethics-framework-glossary-and-methodology, dostęp 14.05.2024]
14. Grubb, S. (2021). How Cybersecurity Really Works: A Hands-on Guide for Total Beginners. No starch press.
15. Guzman, L. & Dyer, S. (2020). Ten questions we're asking about ethics, data, and open source research. Amnesty International [dostępne: https://citizenevidence.org/2020/11/10/ethics-data-open-source/, dostęp 17.05.2024]
16. Hill, M. & Swinhoe, D. (2022). The 15 biggest data breaches of the 21st century. CSO Online [dostępne: https://www.csoonline.com/article/534628/the-biggest-data-breaches-of-the-21st-century.html, dostęp 21.05.2024]
17. Hive Systems (2024). Are Your Passwords in the Green? [dostępne: https://www.hivesystems.com/blog/are-your-passwords-in-the-green?utm_source=tabletext, dostęp 24.05.2024]
18. Kaspersky (n.d.a). How Data Breaches Happen & How to Prevent Data Leaks [dostępne: https://www.kaspersky.com/resource-center/definitions/data-breach, dostęp 21.05.2024]
19. Kaspersky (n.d.b). Top 15 internet safety rules and what not to do online [dostępne: https://www.kaspersky.com/resource-center/preemptive-safety/top-10-preemptive-safety-rules-and-what-not-to-do-online, dostęp 25.05.2024]
20. Kerner, S. M. (2022). Colonial Pipeline hack explained: Everything you need to know. TechTarget [dostępne: https://www.techtarget.com/whatis/feature/Colonial-Pipeline-hack-explained-Everything-you-need-to-know, dostęp 20.05.2024]
21. Kim, D. & Solomon, M. G. (2018). Fundamentals of Information Systems Security, 3rd Edition. Jones & Bartlett Learning.
22. Knight, M. (2021). What Is Data Ethics?. Dataversity [dostępne: https://www.dataversity.net/what-are-data-ethics/, dostęp 14.05.2024]
23. Kosinski, M. (2024). What is a phishing attack? IBM [dostępne: https://www.ibm.com/topics/phishing, dostęp 24.05.2024]

24. McKinsey (2022). Data ethics: What it means and what it takes [dostępne: https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/data-ethics-what-it-means-and-what-it-takes, dostęp 14.05.2024]
25. National Institute of Standards and Technology (NIST) (n.d.). Information security [dostępne: https://csrc.nist.gov/glossary/term/information_security, dostęp 20.05.2024]
26. NSW Government (n.d.). 10 Tips for Cyber Security [dostępne: https://www.digital.nsw.gov.au/sites/default/files/2022-09/top-10-cyber-security-tips.pdf, dostęp 25.05.2024]
27. O'Reilly (2018). Case studies in data ethics [dostępne: https://www.oreilly.com/content/case-studies-in-data-ethics/, dostęp 17.05.2024]
28. PR Newswire (2018). New Survey Finds Deep Consumer Anxiety over Data Privacy and Security [dostępne: https://www.prnewswire.com/news-releases/new-survey-finds-deep-consumer-anxiety-over-data-privacy-and-security-300630067.html, dostęp 20.05.2024]
29. Rubenking, N. J. & Duffy, J. (2023). 12 Simple Things You Can Do to Be More Secure Online. PC mag [dostępne: https://www.pcmag.com/how-to/12-simple-things-you-can-do-to-be-more-secure-online, dostęp 25.05.2024]
30. TechTarget (2024). Top 10 types of information security threats for IT teams [dostępne: https://www.techtarget.com/searchsecurity/feature/Top-10-types-of-information-security-threats-for-IT-teams, dostęp 24.05.2024]

SPIS TABEL

SPIS RYSUNKÓW

Dario Šebalj, dr

Adiunkt na Wydziale Ekonomii i Biznesu w Osijeku, Katedra Metod Ilościowych i Informatyki. Jest wykładowcą w dziedzinie zarządzania procesami biznesowymi, zarządzania projektami ICT, analizy biznesowej i e-logistyki. Opublikował ponad 20 artykułów naukowych w czasopismach i materiałach z międzynarodowych konferencji. W latach 2020-2023 był redaktorem naczelnym czasopisma Ekonomski vjesnik, a w latach 2017-2019 sekretarzem stowarzyszenia Alumni EFOS. Pracował jako badacz w dwóch projektach Erasmus +, a także w projekcie Chorwackiej Fundacji Nauki

ORCID: 0000-0002-8295-7847

Dejan Mirčetić, dr

Pracownik naukowy, Instytut Badań i Rozwoju Sztucznej Inteligencji w Serbii, Adiunkt, Wydział Nauk Technicznych, Uniwersytet w Nowym Sadzie. Dr Dejan Mirčetić uzyskał tytuł doktora w dziedzinie prognozowania popytu i analizy łańcucha dostaw i jest autorem ponad 60 prac naukowych, a jego badania koncentrują się głównie na rozwiązywaniu rzeczywistych wyzwań biznesowych i przemysłowych. W Instytucie Sztucznej Inteligencji w Serbii dr Mirčetić jest odpowiedzialny za projektowanie i rozwój rozwiązań AI dla biznesu i przemysłu. Obecnie kieruje projektem ISIOP, będącym częścią platformy AIPlan4EU w ramach programu badań i innowacji Horyzont 2020. Oprócz pracy w zakresie zarządzania łańcuchem dostaw, dr Mirčetić wnosi znaczący wkład w zastosowanie sztucznej inteligencji w opiece zdrowotnej i przemyśle spożywczym, koncentrując się na innowacyjnych podejściach do optymalizacji procesów i zwiększania wydajności

ORCID: 0000-0002-1602-7908

Michał Adamczak, dr inż.

Pracownik badawczo-dydaktyczny Wyższej Szkoły Logistyki w Poznaniu. Kierownik Katedry Logistyki tej uczeni. Członek Zarządu Polskiego Towarzystwa Logistycznego. Specjalizuje się w tematyce zarządzania, zapasami, zarządzania łańcuchem dostaw i zarządzania produkcją. W swojej pracy wykorzystuje narzędzia analizy danych logistycznych, analizę statystyczną, modelowanie i symulacje procesów, w tym arkusz kalkulacyjny Ms Excel. Autor programów szkoleń otwartych i zamkniętych z wyżej wymienionych tematów. Realizator kilkudziesięciu projektów konsultingowych dla firm handlowych i produkcyjnych. Kierownik naukowiec w wielu projektach realizowanych w ramach programu ERASMUS+. Autor ponad 100 publikacji naukowych w uznanych czasopismach krajowych i zagranicznych. Uczestnik wielu konferencji międzynarodowych.

ORCID: 0000-0003-4183-7264

BUSINESS ANALYTICS SKILLS FOR THE FUTURE-PROOF CHAINS -